Journal of
Experimental Psychology

VOL. 31, NO. 3 SEPTEMBER, 1942

A STUDY OF DISCRIMINATIVE SERIAL ACTION: MANUAL RESPONSE TO COLOR

BY HERBERT B. WEAVER

University of Pennsylvania

INTRODUCTION

The experiment described below was designed to investigate various problems of discriminative serial action, *i.e.* performance characterized by a continuing series of discriminative reactions to a limited number of discrete stimuli. Although this type of serial action has been the subject of numerous experiments over a period of many years, various features of it stand in need of experimental clarification.

Until 1931 studies of discriminative serial action were largely confined to two performance characteristics, speed and accuracy, with scant attention to intra-performance variability. Bills (**2**, **3**, **4**, **5**, **7**, **8**) was the first to study systematically a variability phenomenon which he called 'mental blocking.' This feature of performance, referring to those periods when a subject is unable to respond until a relatively long time has elapsed, had been widely observed since at least 1911 when Woodworth and Wells (**28**) described it in connection with their color-naming test. In a series of experiments Bills studied the laws governing the occurrence of 'blocks,' the factors determining their frequency, and the relation of blocking to fatigue, to refractory phase, to errors, to stammering, and to anoxemia.

The present experiment aims to investigate further some of the questions raised by the work of Bills and of others. Specifically, the problems to be investigated are: (1) the existence of sex differences, differences between the freshman and senior college classes, and individual differences in discriminative serial action measures of speed, accuracy, and variability; (2) the changes in magnitude of measures during the course of performance; (3) the effect of a retest

on performance measures; (4) the relationships among performance measures. Special emphasis is placed on the phenomenon of the occurrence of long response times ('mental blocks').

Apparatus

Detailed description of the apparatus appears elsewhere (27). Essentially it consisted of a black box whose front panel was 22 inches high and 11 inches wide, from the lower portion of which protruded four keys labelled with two sets of four colored rectangles—green, blue, red and yellow from left to right. One set of labels was attached one each directly above each of the keys, the other set directly below a stimulus window in the box. The stimulus window was 5 inches above the keys and was located slightly below eye-level and about one foot in front of the subject as he sat before the box, which rested on a table; the window was one inch in diameter and was formed of ground glass. The labels below the window permitted the subject to fixate the window and also see the labels, the purpose of this arrangement being to eliminate as much as possible the necessity for learning the positions of the keys.

The four stimuli[1] were produced by lamps covered by colored glass screens and arranged in such a way in the box that the stimulus window was lighted uniformly by each; that is, each successive color seemed to come from the same source, so that directional cues with respect to the keys were eliminated. The color screens were so chosen as to produce four unmistakeably different color sensations of equal brightness as subjectively judged; they were not pure primaries, but were easily recognized as green, blue, red and yellow. (Two lamps behind blue filters had to be used to raise the blue illumination to the necessary brightness.)

The four colors appeared an equal number of times in a series of eighty stimuli before the series repeated. The succession of colors was arranged so as to reduce to a minimum the possibility of learning the succession; the order was random except that no color ever followed itself, and the same sequence of four as the order of the keys from left to right or right to left was avoided. Because of this arrangement it is improbable that any learning of the actual sequence of eighty occurred, especially since the subject had no way of knowing when the series began repeating, or, indeed, that there was a fixed series at all. Moreover, the test period was brief, and the instructions to the subject carried a strong suggestion of the impossibility of predicting the next color without making errors. The color-selector mechanism was placed in a separate room because of the noise made by the action of its electromagnet.

The nature of the task presented to the subject by the apparatus is made clear by the instructions to the subject, which are reproduced in full below. The subject was seated in a chair provided with cushions which he could arrange to suit his comfort. The subject's arms were supported by adjustable leather cushions on the table on which the apparatus rested so that his fingers were comfortably placed with respect to the keys in order to reduce muscular fatigue to a minimum.

The recording apparatus was situated across the room and behind the subject. It consisted of a synchronous motor which drove three-hundred-foot rolls of black waxed tape one inch wide at a uniform speed of 103.2 inches per minute. Two heated styli rested on the moving tape, one producing a mark at every right response of the subject, the other a mark at every wrong response.

Subjects

One hundred undergraduate college students were used as subjects. There were 25 freshman males, 25 freshman females, 25 senior males, and 25 senior females. All subjects reported on a voluntary basis. Since one of the principal objects of the experiment was to study sex differences, males and females were matched on the basis of score on the Scholastic Aptitude Test of the College Entrance Examination Board. The mean S.A.T. score in each of the four groups of subjects was 521; the maximum difference in S.A.T. points between matched subjects was 15, and the mean

[1] Four stimuli were chosen for two reasons: (1) in order to control the factor of handedness, an even number was necessary; (2) because of Hansen's (13, p. 337) summarizing statement as follows: "From the experiments of Seashore, Coover, and McComas it is evident that, in tests of continuous discriminative reactions, the most convenient and satisfactory number of movements is four. Less than that number brings into consideration too high a degree of anticipation, while more than four movements has proved to be cumbersome and confusing."

difference was 5.6 (the standard deviation of S.A.T. score distributions is 100). The mean age difference between the freshman males and senior males and between the freshman females and senior females was 33.5 months and 33.9 months respectively. Since the nature of the experimental task required normal color vision, all color-blind and color-weak subjects were eliminated by means of the Ishihara Test. Subjects who said they were unusually fatigued or not feeling well when they reported for the experiment were not tested.

Procedure

All subjects were tested between 3 and 5 P.M. in a windowless ventilated room free of distractions in which constant conditions of illumination were maintained. The only other person present was the experimenter, and he remained in back of the subject during the test performance.

When the subject reported to the laboratory, the experimenter first obtained routine information and gave him opportunity to accustom himself to the laboratory situation. The Ishihara Test was administered. The instructions for the experimental task were then read. These were designed to make the task as clear and simple to the subject as possible. All important details regarding fingering of the keys and relevant features of operation of the apparatus were carefully explained. Opportunity to master these details was afforded by two one-minute practice periods. Emphasis was placed on a 'speed set.' Since the nature of the 'set' is important (Hansen, **13**, p. 337) and since the nature of the task is made clear by the instructions, these are reproduced in full:

(Experimenter sits in position before keys; subject watches.)

"This is a test to see how fast and how accurately you can work. Here are four keys labelled 'green,' 'blue,' 'red' and 'yellow.' Here is a light which can change color; it can be green, blue, red or yellow, the same as the colors of the keys. Here is another label, just like the one above the keys. If you press the key corresponding to the color showing, the light changes right away to another color. The light is now ——, so you press the —— key, and the light instantly changes to another color; now you press the —— key, and it changes again; now ——; now ——; (demonstrate till all four colors have appeared). But you can never tell what color is going to be next; there is no rule or pattern or order in which the light changes. If you press the wrong key, the light doesn't change (demonstrate); or if you press two keys or more than two keys at the same time, the light doesn't change, even though one of them is the right key (demonstrate).

"The test is to see how fast you can make the light change by pressing the right keys. If you press a wrong key, it counts as an error against you; or if you press two keys at the same time, it counts as an error. With that machine over there I'll be able to tell how fast you work and how many errors you make. You're bound to make a few errors, but try to make as few as possible, since errors slow you up.

"The method of fingering the keys is very important. There are a number of things to remember. First of all, you use both hands, two fingers of each hand—index fingers for the inner keys, middle fingers for the outer keys; and you always use the same finger for the same key. Another thing to remember is that the keys go down easily—you don't have to use much force to press them down (demonstrate). The keys must, however, go all the way down. Sometimes, if you just jab at a key this way (demonstrate), it doesn't go all the way down, and the light doesn't change. So what you want is a light but deliberate touch, like this (demonstrate). Once you press a key down and the light changes, you should release it right away. If you continue to hold it down until you're ready to press the next key, you might have two keys down at the same time, and then the light won't change (demonstrate). Do you hear the click back there? If you should ever happen to have two keys down at the same time, that click is a signal to warn you. As a matter of fact, if you should ever accidentally happen to have a wrong key even partly down when you press the right key, the light won't change (demonstrate). But again the click warns you that you're doing it.

"Now I'll let you practice awhile to get used to it. (Subject and experimenter change seats.) Try a few just for touch first. Good. Now keep on for a whole minute (one minute practice; correct any tendency to disobey instructions).

"Good; you do very well. I'm going to let you work about fifteen minutes, without stopping, as fast as you can. When you once start, don't let anything whatsoever stop you or disturb you even for an instant—not that I'm going to let anything do that if I can avoid it. Don't talk or say anything while you work. Don't move around in the chair or move your arms much. Don't look up or look around. Pay no attention to me or to noises from the machine over there, or to

any other noises in here or outside, except the sound of the click to warn you that you have two keys down. Just keep going as fast as you can. Throw every effort into it. And don't look for any particular sequence of colors; the order is completely irregular, and if you try to expect what color is coming next, you'll only make mistakes. The labels to the keys will always be there to guide you, so you don't have to memorize them; you can keep looking at them all the time.

"Now one more minute of practice to be sure you understand everything. (One minute practice.)

"Good. Now make yourself comfortable so you won't have to move. I'm not going to say a word once you start; I'll simply sit over there out of sight. Remember how to press the keys: not too hard, press them all the way down, and only one down at a time. It's easy to forget all this in the course of the work, so try to bear it in mind. Don't be so cautious in an effort to avoid errors that you go very slowly; on the other hand, don't get excited or reckless. Don't talk and don't look around. Be sure not to let anything whatsoever stop you or disturb you even for an instant. Try to keep doing the very best you can the whole fifteen minutes. Make as few mistakes as possible, but above all work at top speed all the time. The most important thing is speed. Is everything clear? I'll tell you when to start."

The experimenter then started the recording apparatus and told the subject to go ahead. During the test period he observed the subject carefully to be sure instructions were obeyed. The experimenter also made notes on qualitative aspects of the performance, and made marks on the recording tape to indicate the end of ten minutes and fifteen minutes. The test performance was defined as the completion of 1000 right responses, thus standardizing the performance in terms of work rather than time. This definition was considered to be preferable to a definition in terms of time, since it was desired to study frequency of occurrence of errors and of long response times; with the amount of work constant, every subject had equal opportunity to make errors and delayed responses, which would not be the case if each subject worked a fixed time at his individual pace. The experimenter had each subject continue until he was sure the subject had completed the required number of responses. Some took fewer than fifteen minutes to do this, others more (the mean time was 14.53 minutes), but no subject was ever stopped before fifteen minutes had elapsed. No subject stopped before he was told to.

The subject was then given a sheet of paper with brief instructions to write a subjective description of his performance. The subjective reports differ widely in content from subject to subject, since an effort was made to avoid suggestion in requesting them; hence they do not lend themselves to group treatment. However, each subject was then asked a series of uniform questions concerning specific aspects of his performance, and the experimenter recorded the replies in the subject's own words; these replies do lend themselves to group treatment.

Finally, each subject was given the Minnesota Vocational Test for Clerical Workers (Number Checking and Name Checking), and was dismissed.

Each of the 25 freshman males was contacted approximately one month after his original performance, and was asked to return for a retest. Only the serial action task was given, the instructions and conditions of performance being identical with the original test. The subject was stopped, however, at the end of ten minutes, since the experimenter limited the retest to only 500 right responses.

Treatment of Individual Results

The reaction time of each of the 1000 right responses obtained from each subject, together with the reaction time of each of the erroneous responses scattered throughout the performance, was determined by measuring each from the record tape with a transparent rule to the nearest sixteenth of an inch. All measures of response time discussed below therefore refer to linear length in sixteenths of an inch, each sixteenth representing a time interval of 0.0363 seconds. (Other constants: one inch = 0.581 seconds; one second = 1.720 inches; one minute = 103.2 inches.)

For each of the 100 subjects each measured response time was recorded in sequence in columns, errors being distinguished from right responses by an appropriate mark in the column. Tables of distribution were then made for the 10 successive hundreds of right responses. The first five tables and the second five tables were combined to form two tables of the halves of the performance, and the halves further combined into a total performance distribution. For each of the resulting 13 distributions the following measures were calculated: mean, median, and a

coefficient of variability obtained by dividing the semi-interquartile range by the median and multiplying by 100.

The number of errors occurring in each of the ten hundreds of right responses, in the first 500, in the second 500, and in the total performance, was counted for each subject. Error times were also tabulated and averaged for the total performance.

One of the purposes of the experiment was to study 'mental blocks' as investigated by Bills and others. Bills defines a mental block as 'a pause in the responses equivalent to the time of two or more average responses' (**2**, p. 231). It was decided, however, that the median response time would be a better criterion of such long response times than the mean, since the median is independent of them whereas the mean is not; that is, the existence of long response times increases the mean response time in direct proportion to their number and length, and thereby renders the mean an unsuitable criterion of their existence; the median, however, remains unaffected. Martinson's (**19**) definition of a mental block as any response time 1.75 or more times the mean response time was rejected on the same grounds as Bills's. Warren and Clark's (**26**) definition of a mental block as a response time equal to or greater than twice the modal response time was also rejected because the mode is a less stable criterion than the median in the present distributions. The exceptionally long response times (referred to hereafter simply as l.r.t.'s) were therefore defined as right response times equal to or greater than twice the median of right response times. Errors were excluded in the definition since it was desired to study l.r.t.'s as a group separate from errors. The number of l.r.t.'s was calculated for each of the tenths, halves, and the total performance of each subject. This was done in two ways: first, the median response time of the total performance was used as a criterion for the number of l.r.t.'s in the total performance; second, the medians of the successive hundreds of response times were used as criteria for the number of l.r.t.'s in their respective hundreds. The number of l.r.t.'s for the total performance as calculated by the first method varies but slightly from the number as calculated by the second, the mean difference amounting to 1.5 l.r.t.'s per subject; this results from the fact that the mean variation of the medians of the tenths of performance from the median of the total performance is only 0.013 seconds. The product-moment correlation coefficient between the number of l.r.t.'s as calculated by the two methods is +.99. In treating data of the total performance, the number of l.r.t.'s as determined by the first method rather than the second was employed, since in the latter case some response times of the total distribution would not meet the criterion of l.r.t.'s but would at the same time be longer than others which did. In treating data of the successive tenths of performance, the number of l.r.t.'s as determined by the second method was employed. In the case of the data on the retest of freshman males, the median response time of the total retest performance of 500 responses was used as a criterion of the number of l.r.t.'s, and this number is compared with the number of l.r.t.'s in the first half of the original test determined by using the median of the first half.

In order to study the comparative performance characteristics of errors and l.r.t.'s, all response times occurring two before, one before, one after, and two after all errors and all l.r.t.'s were located and tabulated for each subject, and the means of the resulting 8 distributions calculated.

Results and Discussion

1. Sex Differences

A sex difference in speed of response in favor of females has frequently been reported for various types of discriminative serial action. For color naming in particular, female superiority has been universally reported. Woodworth and Wells (**28**), using their color naming test, found that adult females took significantly less time on the test, and explained it as probably "not specially related to the color sense, but rather to linguistic facility." Brown (**9**), using an adapted form of Cattell's color naming test, also found that females were significantly faster. Hollingworth (**14**), Lund (**18**), and Telford

(**24**) each obtained similar results, using the same type of test. Using school children in the first to the ninth grades, Ligon (**17**) found girls faster than boys on color naming at each of the 9 grade levels, and proposed that female superiority was to be accounted for by superiority in a 'color naming special factor' rather than in verbal facility as Woodworth and Wells had suggested. Stroop (**23**, p. 658) also found female superiority, and proposed that it may be accounted for on educational and social grounds specifically relating to color: "Education in color is much more intense for girls than for boys, as observing, naming, and discussing colors relative to dress is much more common among girls than among boys." Du Bois (**10**) found females significantly faster than males in naming colors and also in pointing to the appropriate color in a color-key as the response to each color square on the test chart, the degree of superiority being much greater for pointing; he interpreted his findings as corroborating

TABLE 1

GROUP MEANS OF TOTAL PERFORMANCE MEASURES

	Median of Right Response Times*	Mean of Right Response Times*	Number of L.R.T.'s	Number of Errors	Coefficient of Variability†	Number of Right Responses in 15 Minutes
Freshman Males.....	22.60	24.09	16.76	34.44	14.25	1016.4
Freshman Females.....	22.44	24.16	19.40	50.12	14.67	990.5
Senior Males.....	22.01	23.49	15.12	33.32	13.71	1046.8
Senior Females.....	21.12	22.61	16.64	34.96	14.44	1076.5

* In sixteenths of an inch. † $\frac{100\,Q}{\text{Median}}$.

Ligon's hypothesis of female superiority in a color naming special factor, except that he suggested it might better be called a 'color recognizing special factor,' since it appeared to operate more strongly when the response was non-verbal and involved color doubly, *i.e.*, involved both perceiving a color and pointing to its counterpart in the key.

The present experiment sought to throw more light upon the origin of the sex difference in color naming by substituting a purely manual response for a purely verbal one. In this way any linguistic advantage which girls have over boys is removed; but any educational advantage they might have, relating specifically to color, or any advantage traceable to a special color factor, remains. If no sex difference were found, the finding would support the hypothesis that the difference in color naming originates in superior female linguistic facility and would be in conflict with Stroop's educational and social hypothesis as well as Ligon's and Du Bois's special factor hypotheses. If a sex difference were found, the linguistic facility

hypothesis would be difficult to defend, and the evidence would be in harmony with the other hypotheses. The experiment aimed to investigate the existence of a sex difference not only in speed of response but also in accuracy, in intra-performance variability, and in frequency of l.r.t.'s.

Table 1 shows the means of the various measures of the total performance for the four groups of subjects. The median of right response times measures the 'typical' reaction time of each group uninfluenced by l.r.t.'s. The mean of right response times includes the influence of l.r.t.'s. The number of right responses in fifteen minutes is an over-all measure of efficiency of performance, including the influence of errors.

Table 2 shows the results of application of the *t*-test [2] for the significance of the differences between the means of Table 1; the

TABLE 2

t-TESTS OF SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF TOTAL PERFORMANCE MEASURES FOR THE SEXES

	Median of Right Response Times	Mean of Right Response Times	Number of L.R.T.'s	Number of Errors	Coefficient of Variability	Number of Right Responses in 15 Minutes
Freshmen						
t	0.244	0.095	0.578	1.658	0.393	0.805
P	.809	.900+	.569	.110	.697	.429
Direction	Females shorter	Males shorter	Females greater	Females greater	Females higher	Males greater
Seniors						
t	1.259	1.429	0.353	0.196	0.709	0.743
P	.220	.169	.727	.847	.485	.465
Direction	Females shorter	Females shorter	Females greater	Females greater	Females higher	Females greater

significance of the sex differences has been tested at the freshman level by comparing the freshman male mean for each measure with the freshman female mean, and similarly at the senior level. The probability values corresponding to each *t* show that all of the obtained differences may have arisen through chance. Regardless of what performance measure is used, no significant sex difference exists.

The sex difference on the color naming test therefore depends on the fact that the response is verbal and not on the fact that the stimulus is color, since it disappears when the response is manual. This is in accord with the almost universally observed superiority of

[2] Because of the low correlations between the Scholastic Aptitude Test scores and the performance measures, the form of the *t*-test in which differences between matched subjects are studied was not used, and the group comparison form was instead employed.

females in verbal functions. The finding is in conflict with Stroop's, Ligon's, and Du Bois's hypotheses. It is in accord with Hansen's finding (**13**) that there is no sex difference in speed or accuracy of response when the task consists of pressing one of four keys in response to spaced auditory stimuli. It is in possible conflict, however, with Du Bois's finding, referred to above, of significant female superiority when the response consists of pointing instead of naming. It may also be in conflict with the report (**15**) of female superiority on the serial discrimeter, which requires pressing keys in response to digits serially exposed.

The data were tested for a sex difference in group variability by calculating the critical ratio of the difference between the standard deviations of the group distributions for each performance measure at both the freshman and senior levels. The standard deviations tested are shown in Table 4. Of the 10 critical ratios, all are below 2.00 except that for number of l.r.t.'s at the senior level, which is 2.62 in the direction of greater group variability for males.

2. College Class Differences

In order to determine whether college class is an important variable in serial action experiments of the present type, comparisons of the classes were made in a manner similar to comparisons of the sexes as indicated in the foregoing section. The significance of class differences was thus tested for both sexes by comparing freshman males with senior males and freshman females with senior females. Table 3 shows the *t*-values and their probability values obtained for the differences between the means in Table 1. None of the values for males is significant, but for females the differences in the three

TABLE 3

t-TESTS OF SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF TOTAL PERFORMANCE MEASURES FOR THE CLASSES

	Median of Right Response Times	Mean of Right Response Times	Number of L.R.T.'s	Number of Errors	Coefficient of Variability	Number of Right Responses in 15 Minutes
Males						
t	0.800	0.680	0.318	0.195	0.475	0.796
P	.432	.502	.753	.847	.639	.434
Direction	Seniors shorter	Seniors shorter	Freshmen greater	Freshmen greater	Freshmen higher	Seniors greater
Females						
t	2.100	2.549	0.773	1.691	0.242	2.845
P	.046	.017	.447	.102	.811	.01–
Direction	Seniors shorter	Seniors shorter	Freshmen greater	Freshmen greater	Freshmen higher	Seniors greater

measures of speed approach significance in favor of seniors, that for median response time being of possible significance, that for mean response time being probably significant, and that for number of right responses in 15 minutes being the most significant of the three.

In order to test whether the superiority of the senior girls over the freshman girls in the number of right responses in 15 minutes is traceable to the age difference between the two groups, the following calculation was made: The mean number of right responses in 15 minutes was determined for the oldest 9 subjects and the youngest 9 subjects in each of the two groups of 25. For the freshman girls, where the mean age difference between the two groups of 9 is 11.5 months, the mean number of right responses in 15 minutes for the younger group is 988.4, and for the older group, 983.1. For the senior girls, where the mean age difference between the two groups of 9 is 15.1 months, the mean for the younger group is 1103.7, and for the older group, 1068. In neither case is the difference significant. Although the age difference in both cases is slight, the fact that the difference in the performance measure is numerically in favor of the younger group in both cases indicates that age is not the factor responsible for the significant difference in favor of the senior class. The actual origin of the difference cannot be determined on the basis of the present data. Since it is improbable that the difference arises out of difference in scholastic attainment between the two classes (see 'Individual Differences,' below), the origin is probably to be sought in difference in adjustment to the laboratory situation between the two groups.

Class difference in group variability was tested in the same manner as sex difference in group variability described in the foregoing section. Of the ten critical ratios, all were below 1.50.

3. *Individual Differences*

By far the greatest source of variation in the data, for all performance measures, is that traceable to individual differences. Their magnitude is shown by the maxima, minima, and standard deviations of the group distributions of total performance measures in Table 4. The table shows for each performance measure and for each group of 25 subjects, the maximum measure made by any one subject, the minimum measure made by any one subject, and the standard deviation of the distribution of the measures of the subjects around their group mean. That the individual differences are highly reliable and represent a stable 'personal idiom' of performance is shown in the case of the freshman males by product-moment correlation coefficients between measures of the first half of the original performance and the retest performance (of identical length) one month

TABLE 4

MAXIMA, MINIMA, AND STANDARD DEVIATIONS OF GROUP DISTRIBUTIONS OF TOTAL PERFORMANCE MEASURES

	Mean of Right Response Times*			Number of Errors			Number of L.R.T.'s		
	Maximum	Minimum	Sigma	Maximum	Minimum	Sigma	Maximum	Minimum	Sigma
Freshman Males	29.930	19.109	2.976	153	4	34.887	85	2	17.729
Freshman Females	29.164	20.615	2.156	138	7	31.921	56	0	14.367
Senior Males	31.584	18.190	3.264	131	5	29.097	85	0	18.742
Senior Females	26.927	18.235	2.177	148	8	29.996	44	4	10.595

	Coefficient of Variability†			Number of Right Responses in 15 Minutes		
	Maximum	Minimum	Sigma	Maximum	Minimum	Sigma
Freshman Males	24.82	5.01	5.339	1286	748	129.27
Freshman Females	21.55	9.07	3.584	1152	807	96.67
Senior Males	26.74	8.42	4.090	1328	706	145.71
Senior Females	20.25	8.79	3.123	1351	904	116.17

* In sixteenths of an inch. † $\frac{100\,Q}{\text{Median}}$.

later. For mean of right response times, the coefficient is + .907; for number of errors, + .861; for number of l.r.t.'s, + .870; and for coefficient of variability, + .896 (from Fisher's Table V.A. (11), a coefficient of 0.5069 indicates a significant correlation at the .01 probability level for the given sample size).

Individuals extreme in either direction in speed, accuracy, variability, and frequency of l.r.t.'s are of particular interest. Since continuous discriminative responding is basic to many job operations, determination of an individual's performance characteristics might serve as a useful index to his job competency. Although the serial discrimeter has been employed for this purpose, its use is restricted to measurement of speed and accuracy, whereas variability and frequency of l.r.t.'s are seen to constitute additional performance characteristics knowledge of which might be desirable.

In an effort to investigate the origin of individual differences, the performance measures were correlated with Scholastic Aptitude Test score and with scores on the Minnesota Vocational Test for

Clerical Workers, which has been referred to as a test related to discriminative serial action. Table 5 shows the resulting coefficients. There is clearly none but a chance relation between S.A.T. score and any of the performance measures; the source of individual differences therefore does not lie in the sort of mental ability measured by the Scholastic Aptitude Test. Both the Minnesota Number-Checking and Name-Checking scores bear a slight but reliable negative relation to the mean of right response times, indicating that individuals who are fast in 'checking' tend to be fast also in serial response. Although the correlations between the Minnesota scores and the other performance measures fall short of significance, they are consistent in the direction of an inverse relation between scores and frequency of errors, frequency of l.r.t.'s, and coefficient of variability. The number of errors on the Minnesota tests and the number of errors on the serial action test show no relation. Determination of the

TABLE 5

PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN PERFORMANCE MEASURES AND SCHOLASTIC APTITUDE TEST SCORE AND MINNESOTA VOCATIONAL TEST FOR CLERICAL WORKERS (NUMBER-CHECKING AND NAME-CHECKING)

	Mean of Right Response Times	Number of Errors	Number of L.R.T.'s	Coefficient of Variability		Number of Errors
Scholastic Aptitude Test Score	−.121	+.078	−.048	−.052	Number of Errors on Minn. Clerical Number-Checking	+.034
Minn. Clerical Number-Checking Score	−.346	−.177	−.134	−.137	Number of Errors on Minn. Clerical Name-Checking	+.038
Minn. Clerical Name-Checking Score	−.313	−.157	−.199	−.190		

From Fisher's Table V.A. (11, p. 214), the following r's are significant at the indicated probability levels for the given sample size, and can be used to judge the significance of the r's in the above table:

for $r = .1682$, $P = .10$
for $r = .1998$, $P = .05$
for $r = .2362$, $P = .02$
for $r = .2607$, $P = .01$.

The coefficients in this table and in Table 8 were obtained in the following manner: Since four different subject groups had been used, and since the correlation between any two variables might operate in a differential way from one group to another, r's were calculated separately for each of the four groups in each case. In order to test for the possibility of differential correlation, Fisher's z-transformation was applied to each of the four r's, and the difference between the highest and lowest z was tested by its standard error to ascertain whether the two groups represented by the two z's could have been drawn from unequally correlated populations. Since in every case no evidence was found for concluding that the groups were not drawn from equally correlated populations, the four r's of the groups were combined into the single value shown in the table, again by the use of the z-transformation as described by Fisher (11).

origin of individual differences thus remains for further investigation. Until they are accounted for and controlled, precise and efficient experimentation is difficult because of the great variability they superimpose on experimental variables.

4. Changes in Magnitude of Performance Measures During the Course of Performance

Figures 1, 2, and 3 show the changes in magnitude of performance measures when the performance is divided into successive tenths

Graphs Showing Changes in Magnitude of Performance Measures During Course of Performance

Figure 1

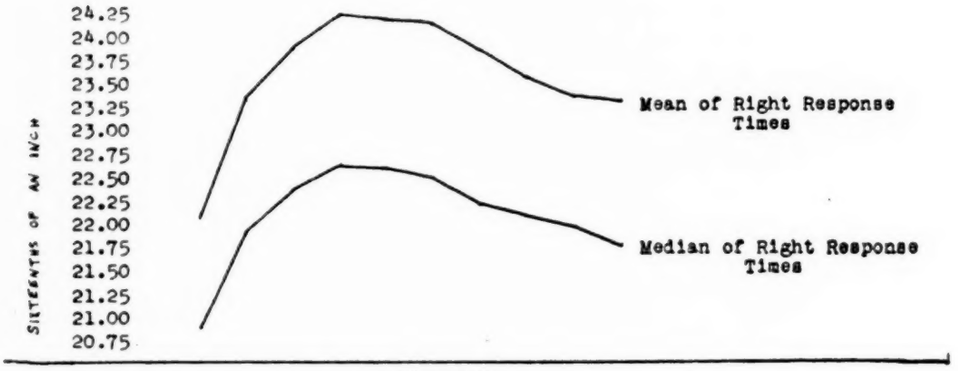


Figure 2

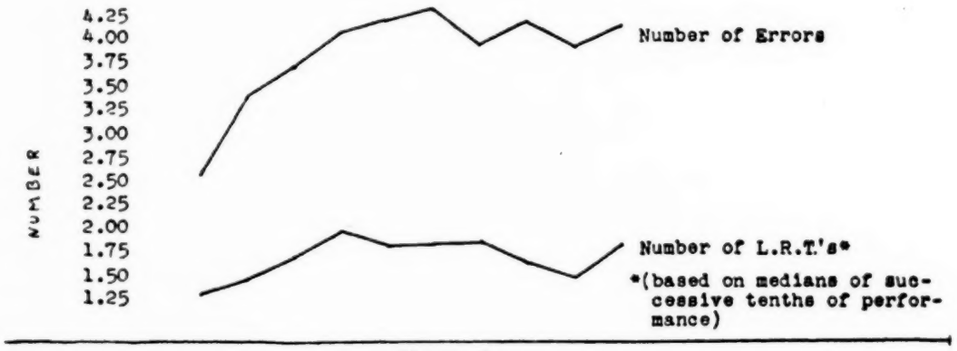
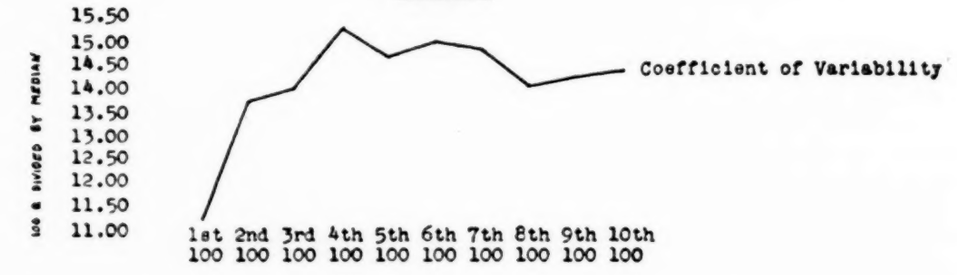


Figure 3



and the results for all subjects are combined. The curves have a common characteristic of showing an initial decrement of performance extending approximately to the mid-point: the mean and median of

response times increase, the number of errors increases, the number of l.r.t.'s increases, and the coefficient of variability increases. The curves have the additional feature of showing a recovery in the second half of the performance, with the possible exception of the curves of errors and of l.r.t.'s.

The nature of the factor or factors which account for the initial decrement and subsequent recovery cannot be determined from the present data. The curves have some resemblance to a monotony curve, but the recovery period is too prolonged to be considered an end-spurt. The curves have none of the characteristics of a fatigue curve. Bills (**2**) has presented evidence that increase in frequency of mental blocks is a sensitive index of fatigue. The fact that the l.r.t.'s curve does not rise in the latter half of the performance indicates that fatigue was not a dominant factor if l.r.t.'s are a fatigue phenomenon. Moreover, when the number of l.r.t.'s in the first half and in the second half of the performance are compared for each subject, it is found that 46 subjects had fewer in the second half, 43 had more, and 11 had the same number.

5. *Effect of Retest on Performance Measures*

The comparative magnitude of the performance measures of the first half of the original test for the freshman males and of a retest given the freshman males under identical conditions one month later may be seen from Table 6. Marked improvement has occurred in

TABLE 6

MEANS OF PERFORMANCE MEASURES OF TEST (FIRST HALF) AND OF RETEST FOR FRESHMAN MALES

	Median of Right Response Times*	Mean of Right Response Times*	Number of L.R.T.'s	Number of Errors	Coefficient of Variability†	Number of Right Responses in 10 Minutes
Test	22.65	24.06	7.84	15.88	13.87	671
Retest	20.07	20.91	3.52	15.96	11.63	775

* In sixteenths of an inch.

† 100 Q divided by median.

every respect except accuracy. That the improvements are significant is obvious from the results of application of the t-test (individual comparison form) to the differences between the measures of test and retest in Table 6; the t-values and their probability values are shown in Table 7. Further facts include the following: In the case of every subject, without exception, the mean of right response times decreased on the retest, the median of right response times decreased, the number of right responses in ten minutes increased,

TABLE 7

t-TESTS OF SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF PERFORMANCE MEASURES OF TEST AND OF RETEST

	Mean of Right Response Times	Number of L.R.T.'s	Number of Errors	Coefficient of Variability
t	6.614	5.812	0.038	6.606
P	.01−	.01−	.900+	.01−
Direction	Retest shorter	Retest smaller	Retest greater	Retest lower

and the coefficient of variability decreased. Of the 25 subjects, 23 had fewer l.r.t.'s, one had more, and one had the same number on the retest. One subject had no l.r.t.'s on the test, but 9 had none on the retest.

Two points of interest appear in the above findings. First, the degree and consistency of improvement in speed, variability, and number of l.r.t.'s are surprising in view of the facts that one month elapsed between the test and the retest, and the subjects had had only a little more than fifteen minutes of experience with the task before the retest. The explanation of the improvement probably lies not only in the persistence of specific practice effects on the task but also in habituation to the laboratory situation at the retest. The second point of interest is the fact that accuracy showed no change from test to retest: there is a differential effect between accuracy and the other performance characteristics; reference will be made to this below.

6. Interrelations of Performance Measures

In order to study the interrelations of the performance measures, product-moment correlation coefficients between measures were calculated; these are shown in Table 8. As the table shows, speed as measured by the mean of right response times bears no relation to accuracy—faster individuals are neither less accurate nor more accurate than slower individuals; this agrees with Hansen's finding (**13**). The number of errors and the number of l.r.t.'s bear a slight but reliable positive relation to each other, showing a tendency for individuals who have many l.r.t.'s to make many errors. The coefficient of variability shows a similar slight but reliable relation to both the mean of right response times and the number of errors, indicating that variability tends to be inversely related to speed and to accuracy. A high positive relation obtains between the coefficient of variability and the number of l.r.t.'s. These measures are statistically independent; the fact that they correlate so highly indicates that the same factor which is responsible for the variability of the

fifty percent of response times around the median in the individual subject's distribution of response times is responsible also for the frequency of those response times which depart most widely from the median.

For determining the relation of speed to the frequency of l.r.t.'s, it was necessary to use the median rather than the mean of right response times as the measure of speed in order to avoid a statistical artifact, since a large number of l.r.t.'s would inflate the mean and yield a spurious coefficient. The obtained coefficient shows an absence of relation between speed and the frequency of l.r.t.'s. When the mean of right response times and the number of l.r.t.'s

TABLE 8

PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN MEASURES OF TOTAL PERFORMANCE

	Mean of Right Response Times	Number of Errors	Number of L.R.T.'s
Number of Errors	−.037	—	+.312
Coefficient of Variability	+.362	+.291	+.817
Median of Right Response Times	—	—	+.062

Coefficients in the above table were calculated in the same manner as those in Table 5, and their significance can be judged on the same basis as that indicated for Table 5:

for $r = .1682$, $P = .10$
for $r = .1998$, $P = .05$
for $r = .2362$, $P = .02$
for $r = .2607$, $P = .01$.

are correlated, the coefficient is + .300, practically the same[3] as that obtained by Bills (**2**) on a color-naming test. Bills obtained a coefficient of − .33 between speed of response as measured by the number of responses per minute and the frequency of mental blocks, and a coefficient of − .70 between speed of response and the length of blocks, which led him to conclude (p. 236) that "the same factor which is responsible for rapid mental functioning tends also to minimize the need for the enforced rests or blocks," and (p. 244) that "Individuals who respond rapidly tend to have less and shorter blocks than slow individuals." These conclusions must be rejected as being based on arithmetical rather than on psychological relations for the following reasons: The larger the number of blocks, the lesser must be the speed as measured by the number of responses in a given time; hence the inverse correlation obtained by Bills between speed of response and the frequency of blocks is spurious. Regarding the coefficient of − .70 between speed of response and the length of

[3] Since the mean of right response times in the present data is really a measure of slowness rather than speed, the sign of the coefficient must be reversed if interpretation is made directly in terms of speed.

blocks, the fewer the responses in a given time, the higher the mean of response times and hence the higher the mean of those response times above twice the mean response time: this coefficient also is therefore spurious.

7. *Distribution of Response Times of Individual Subjects*

Examination of the distribution of response times of the total performance of individual subjects shows that the distributions are characterized by a high degree of concentration around the median, but that there is a pronounced skewness in the direction of longer response times. The high degree of concentration around the median is evident from the fact that the mean coefficient of variability for all subjects combined is 14.27 (percent of the median), representing a time range of 0.114 seconds; thus, on the average, fifty percent of a subject's response times vary from his median response time by only 114 milliseconds. (When the absolute measure, semi-interquartile range, instead of the relative measure, coefficient of variability, is used, the figure is 115 milliseconds.)

Figure 4 shows graphically the distributions of the right response times of four subjects selected at random as a sample from the total group of subjects. In each case there are few short response times, the major group rising abruptly to a high frequency level, then descending less abruptly and levelling off with many irregularities to form a long tail.

Bills (**5**, p. 562) has reported that "for such simple tasks as color-naming, form-naming, code substitution, addition, naming opposites, and the like . . . if a given subject's reaction times are plotted in the form of a frequency polygon, in which the base represents reaction time in eighths of a second, it is found that two distinct modes appear, one of which is due to the central tendency in normal reaction times, while the other shows the central tendency of the blocks. The block mode appears at a point representing approximately double the normal reaction time mode." Again Bills (**5**, p. 564) says: "Frequency polygons were constructed in which the frequency of length of reaction times was determined. This brought out two distinct modes, one for normal reaction times and the other for blocks. These stood out distinctly with practically no overlapping, thus demonstrating unquestionably the reality of the block." The distributions of all of the individual subjects were inspected for bimodality, but no indications of two distinct modes were found. Graphical representation, as in Fig. 4, also fails to disclose clear bimodality. The appearance of the distributions is thus in apparent conflict with Bills's report and interpretation. It is to be noted,

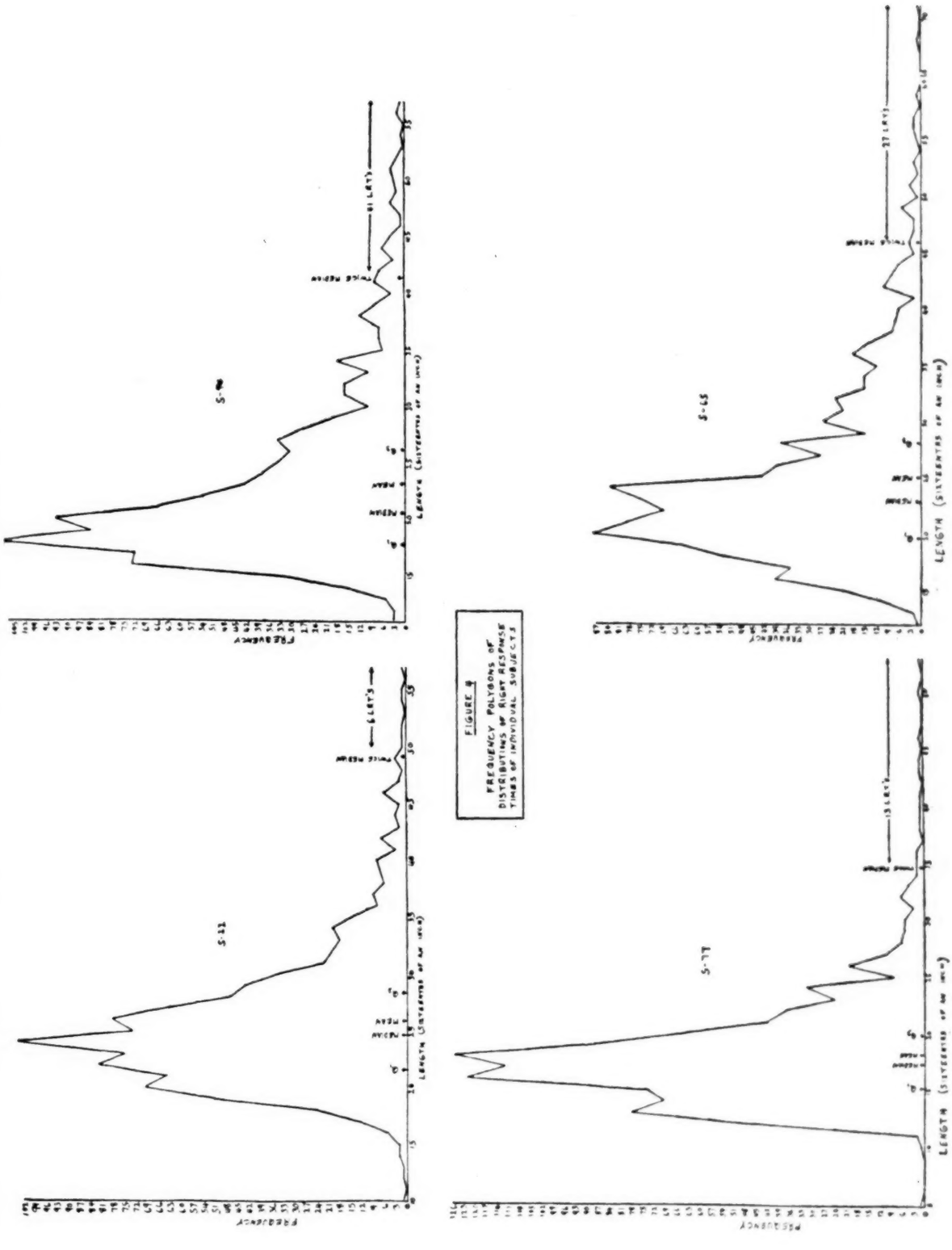


FIGURE 8
FREQUENCY POLYGONS OF DISTRIBUTIONS OF RIGHT RESPONSE TIMES OF INDIVIDUAL SUBJECTS

however, that the number of l.r.t.'s in the present data constitutes a very small proportion of the total number of response times (the mean number is 1.7 percent of the total, far smaller than the number of mental blocks in Bills's data), and could not be expected to produce a clearly discernible mode. The appearance of the distributions is therefore not inconsistent with that of distributions made up of a very large sample of one population of response times and a very small sample of another.

8. Subjective Reports on L.R.T.'s

Investigators of discriminative serial action have frequently reported the phenomenon quantified by Bills as 'mental blocking,' and noted that their subjects were often spontaneously voluble on the obstructions they had experienced. Thus, Woodworth and Wells in 1911 (**28**, p. 52) wrote: "The behavior of the subjects during the color-naming test shows periods of hesitation and obstruction, and even of false reaction—rather a strange phenomenon, in view of the great familiarity of the names and their correct use immediately before. The subject is aware of this inhibition, and it is a strange experience for him." Garrett and Lemmon in 1924 (**12**, p. 432) wrote: "One of the most striking features of the color-naming test is the inhibition or interference which the subject experiences during its performance. He starts out confidently, naming the colors rapidly, but before he has gone very far he begins to hesitate and make mistakes; sometimes he will even stare at a color for several seconds before he can give its name. There are great individual differences in the amount of interference experienced; some subjects feel practically none, some experience a little but by making an effort overcome it, while others are bothered throughout the entire performance." Jersild in 1926 (**16**, p. 34) wrote: "Those who have taken the color and form-naming tests will have noticed an unpleasant feeling of what might be called 'thwarted' mental activity in reading off the successive items. The verbal responses seem to lag behind the immediate perceptual process."

It was decided to obtain subjective reports on the phenomenon as it appears in manual response in order to get both a subjective description and an objective measure. Among the questions asked the subject after his performance were the following:

Did you notice that you hesitated or stopped for a moment every once in awhile?

Every one of the 100 subjects answered 'yes' to this question.

What did this feel like when it happened? Can you describe it?

Twenty-one subjects declared themselves unable to give a verbal description. The remaining 79 gave varied answers, the following being samples:

"I'd lose control."

"A feeling of confusion."

"It would be like a stutter."

"My mind went blank."

"A feeling of disgust."
"I was just sitting there."
"My fingers didn't respond."
"My hands just stopped working."
"I'd see the color and not have the faintest idea what key to press."
"I just stopped."
"I couldn't do anything."
"A resting period."
"Sort of a block."
"A feeling of hypnotism."
"Felt like a snag."
"I just had to stop—nothing would happen."
"Something would happen—I don't know what it was."
"Sort of baffled feeling."
"A surprise—I just stopped."
"I saw the color but just couldn't push the key."
"All of a sudden I'd lose connection between colors and keys—I'd lose control—it was disconcerting."
"There was a definite break—my mind was a blank—I didn't know what to hit."
"I was confused and annoyed, but there wasn't much I could do about it."

What went on in your mind when this happened? What were your thoughts?

Thirty-seven subjects reported their mind was a blank, 15 that their thoughts were confused, 11 that they were thinking of 'Nothing in particular,' 17 that they were thinking of the colors or the keys, and 20 were unable to give a definite answer.

Was the trouble in your fingers or was it mental?

Twenty-one subjects reported that the trouble was in their fingers, 54 that it was mental, 11 that it was both mental and in their fingers, and 14 could not give a definite answer.

When it happened would you say you felt as though you were 'blocked'?

Seventy-six answered 'yes,' 9 answered 'no,' and 15 could give no definite answer.

It is clear from the foregoing that the 'hesitations' experienced by the subjects were beyond their voluntary control, and the data support Bills's statement that blocks are automatic or non-voluntary pauses. A point of interest is the unpleasant affective tone associated with the hesitations in the case of many of the subjects. This unpleasantness is undoubtedly not inherent in the momentary cessation of responding itself. It probably arises out of the feeling of frustration induced by the speed 'set' given in the instructions; the delay in response may produce in the subject a feeling of 'failure' to meet his level of aspiration based on preceding response times.

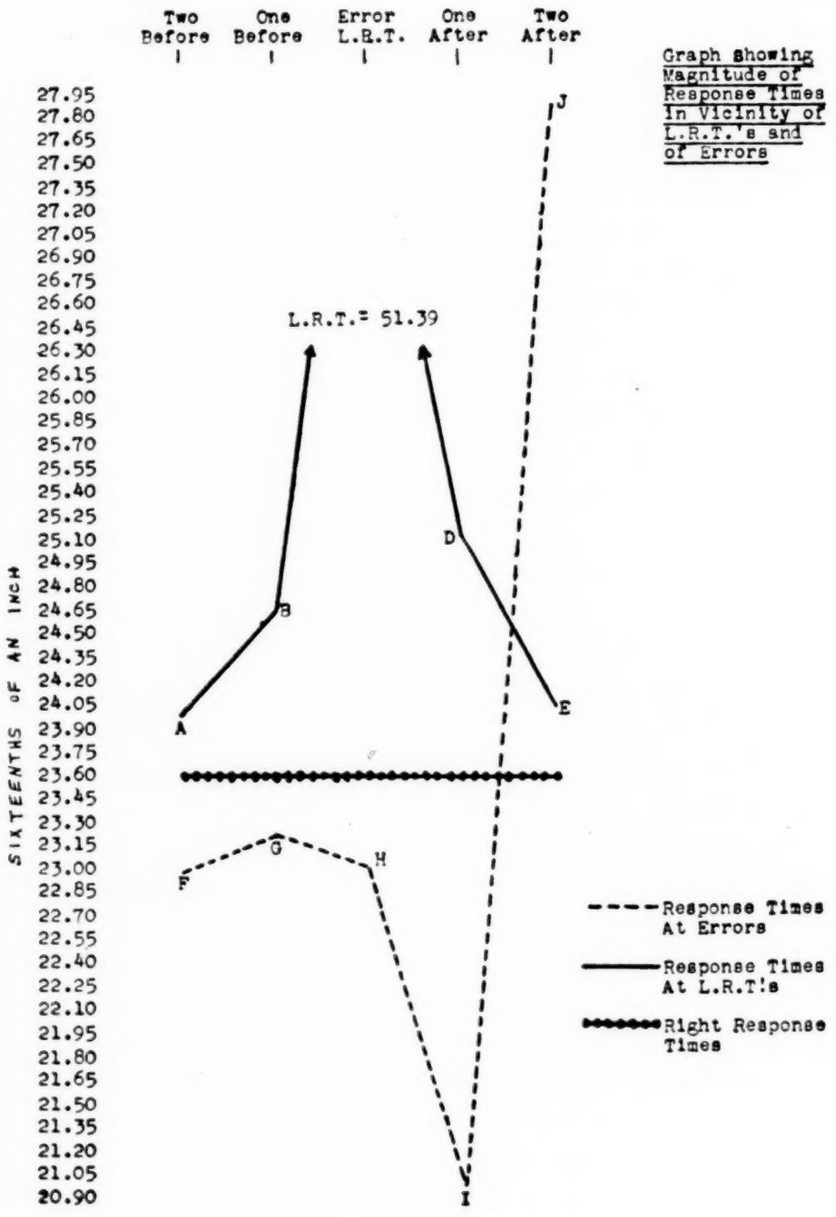
It can be concluded from the subjective reports that most subjects have a qualitatively distinct experience in the course of performance which can be appropriately called 'mental blocking.' The extent to which the objective measurement of long response times corresponds to the subjective experience of mental blocks is uncertain, however, and could be determined only by a specially designed experiment. That the subjective experience and the objective measure do not completely coincide is revealed by the fact that although all of the 100 subjects reported that "they hesitated or stopped for a moment every once in awhile," 3 had no l.r.t.'s, and 2 of the 3 reported that when the hesitations occurred they felt as though they were 'blocked.' Objective measurement at present, involving as it does an arbitrary

and quantitative criterion, thus does not yield a completely valid index of the subjective experience. Until an objective measure can be made which can be demonstrated to coincide with the subjective experience, it is safer to restrict the term 'mental blocks' to the latter and to refer to the former merely as a measure of long response times.

9. *Performance Characteristics of L.R.T.'s and Errors*

L.r.t.'s and errors are occurrences of special interest in the course of serial action. Bills has advanced the hypothesis that they are related in that both result from a low condition of neural functioning

FIGURE 5



(2). In order to study their comparative and individual performance characteristics, all response times in the immediate vicinity of each (two before to two after) were averaged separately for each position for each subject, and the combined averages for all subjects studied graphically and statistically. Figure 5 shows the relative and absolute magnitude of response times in the regions studied, as well as their magnitude with respect to the mean of all right response times. (Separate graphs for the four subject groups show identical relations and are not reproduced.) Table 9 shows the results of

TABLE 9

t-TESTS OF SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF RESPONSE TIMES AT CORRESPONDING POINTS IN THE VICINITY OF ERRORS AND OF L.R.T.'s

	Two Before	One Before	One After	Two After
t.....	4.236	4.729	7.995	7.998
P.....	.01—	.01—	.01—	.01—

application of the *t*-test for the significance of the difference between the means of response times at corresponding positions in the vicinity of errors and of l.r.t.'s (in terms of the graphical representation in Fig. 5, these tests are between points *A* and *F*, *B* and *G*, *D* and *I*, and *E* and *J*). Table 10 shows the results of application of the

TABLE 10

t-TESTS OF SIGNIFICANCE OF DIFFERENCE BETWEEN MEAN OF RIGHT RESPONSE TIMES AND MEANS OF RESPONSE TIMES IN THE VICINITY OF ERRORS AND OF L.R.T.'s

	Mean of Right Response Times and Mean of Response Times								
	Two Before an L.R.T.	One Before an L.R.T.	One After an L.R.T.	Two After an L.R.T.	Two Before an Error	One Before an Error	Of Errors	One After an Error	Two After an Error
t.....	2.577	3.878	5.595	2.239	5.248	2.848	3.381	6.153	8.709
P.....	.012	.01—	.01—	.0295	.01—	.01—	.01—	.01—	.01—

t-test to the difference between the mean of right response times and the mean of response times at each position in the vicinity of errors and of l.r.t.'s as well as the mean of error times (in terms of the graph, these tests are between each of the lettered points in turn and the horizontal line).

All of the differences tested in Table 9 reach a high level of significance, showing that the pattern of response times in the vicinity of errors differs reliably from that in the vicinity of l.r.t.'s. Response times two before, one before, and one after an l.r.t. are significantly longer than corresponding response times near errors,

while those two after an l.r.t. are significantly shorter than those two after an error. L.r.t.'s and errors thus have qualitatively different performance characteristics.

Considering Table 10 and Fig. 5, it is seen that response times two before an l.r.t., one before an l.r.t., and one after an l.r.t. are significantly longer than the mean of all right response times, and those two after an l.r.t. are probably significantly longer but at a lower level of significance; those two before an error, one before an error, the error time itself, and those one after an error are significantly shorter than the mean of all right response times, while those two after an error are significantly longer. L.r.t.'s are thus both preceded and followed by response times which behave in a characteristic manner. It can therefore be concluded that l.r.t.'s represent a qualitatively distinct population of response times. The fact of longer response times preceding l.r.t.'s is in harmony with Bills's hypothesis that blocks result from a low condition of functioning. However, if, as Bills also contends, they are automatic rest pauses, they are rests not immediately followed by a facilitation of responding but by a period characterized by response times longer than the average. The explanation of the fact that response times do not immediately return to the mean level after an l.r.t. (or go below it, if l.r.t.'s are rest pauses) may be the inhibitory effect of the unpleasant affective tone which subjects report is associated with 'hesitations.'

The data regarding errors require a different interpretation from that for l.r.t.'s. The fact of the presence of shorter than average response times before an error points away from Bills's interpretation of errors as resulting, like l.r.t.'s, from a low condition of neural functioning. The data indicate, instead, that a subject makes an error because he is exceeding his natural rate of response. This is corroborated by the testimony of many subjects, when asked why they made errors, that errors occurred because they tried to go too fast. It may even be conjectured that the slight but reliable positive correlation between the frequency of errors and the frequency of l.r.t.'s reported in a previous section results from the circumstance that those subjects who try most persistently to speed up their responses beyond their natural rate, and hence who make the most errors, are the subjects who are most acutely aware of the need for speed because of their many long response times; this, of course, is conjecture, but serves to indicate the possibility of a purely psychological explanation of the correlation between errors and l.r.t.'s instead of leaving the matter as requiring a neurological interpretation. The greatly reduced response time immediately following an error may be accounted for by the subject's immediate realization of

the incorrectness of the response even during the act of pressing the wrong key but too late for inhibition of pressure on the key, an experience common to typists; moreover, following the pressing of the wrong key, the subject had a choice of one less key as the correct one, which would decrease the element of choice in the task and make for faster response. The increased response time immediately following the correction of an error may result from inhibition of response consequent upon awareness of error, again an experience common to typists.

The above interpretation of errors and of l.r.t.'s is in harmony with the finding reported in a previous section that l.r.t.'s decreased significantly in number upon retest, whereas errors remained the same in number. The tendency to exceed the natural rate of response because of instructions to work as fast as possible was the same for test and retest and resulted in equal numbers of errors, whereas the previous specific practice on the task and habituation to the laboratory situation at the retest were reflected in fewer periods of lowered functioning and hence fewer l.r.t.'s.

10. Frequency of L.R.T.'s and of Errors

Various investigators of serial action have reported widely differing figures for the frequency of occurrence of long response times and of errors. Consideration of all the reports shows that there are wide differences in the frequency of long response times depending on the exact conditions involved. Even when conditions are closely comparable, considerable discrepancies appear. For example, Bills (**3**, **4**) reported two studies each of which involved automatic kymographic recording of vocal response to the five-color Woodworth-Wells Chart, presumably with similar or identical instructions and with identical definitions of long response times; in one case long response times occurred on the average once in 17 seconds, in the other case once in 9 seconds, a ratio of almost two to one; the smaller frequency, moreover, occurred in the case of the longer work period, which is the opposite of what is to be expected from Bills's conclusion that blocks are a fatigue phenomenon. Again, Bills (**2**) and Warren and Clark (**26**) reported on vocal response to a six-color chart, using the same method of recording and the same type of instructions to subjects; although in one case the mean response time and in the other the modal response time was used as a criterion of long response times, Bills found a mean frequency of one long response time in 17 seconds for a ten-minute work period, while Warren and Clark found frequencies as high as one in 2.9 seconds for a much shorter work period (which presumably should make for a smaller number), representing a ratio of almost 6 to 1. Although Warren and Clark

used but 4 subjects in their experiment, another group of 4 subjects showed a closely similar frequency of one long response time in 3.4 seconds. Apparently, then, even slight variations in conditions cause wide differences in the frequency of long response times. The data of the present experiment show a much lower frequency than any previously reported despite the fact that the median response time, which is always shorter than the mean response time, was used as a criterion.

Unfortunately, reports on the frequency of errors are scanty or absent. The data of the present experiment, however, show a frequency of errors much more closely approximating those reported than is the case for long response times.

The wide variations observable in the reports demonstrate that much further study of the psychological conditions which determine the occurrence of long response times is needed before the phenomenon can be regarded as placed on a sound theoretical basis; until it is placed on a sound basis, construction of a neurological theory is premature.

Summary and Conclusions

One hundred college adult subjects were tested on a discriminative serial action task consisting of manual response to color; their response times were automatically recorded and measured. The performance was analyzed for speed, accuracy, and variability; special attention was given to long response times, called by Bills and others 'mental blocks.'

Results are reported on sex differences, college class differences, individual differences, changes in performance measures during performance and at a retest, interrelations of performance measures, frequency distributions of response times, subjective reports, performance characteristics of long response times and errors, and frequency of long response times and errors in comparison with related studies.

The following conclusions are drawn:

1. No sex difference exists in any performance measure.
2. All performance measures are highly reliable.
3. Performance is independent of scholastic aptitude.
4. All performance measures except accuracy improve upon retest.
5. Speed of response is unrelated to accuracy or to frequency of long response times; variability is closely related to frequency of long response times.
6. Objective measurement of long response times is not a completely valid index of the subjective experience of mental blocking.

7. Long response times and errors have qualitatively different performance characteristics.
8. Long response times are a qualitatively distinct phenomenon.
9. Errors occur as a result of responding too rapidly.
10. Additional study of the phenomenon of long response times is needed before it can be regarded as placed on a sound theoretical basis.

(Manuscript received December 24, 1941)

Bibliography

1. Berdie, R. F. Effect of benzedrine sulphate on blocking in color naming. *J. exp. Psychol.*, 1940, **27**, 325–332.
2. Bills, A. G. Blocking: A new principle of mental fatigue. *Amer. J. Psychol.*, 1931, **43**, 230–245.
3. Bills, A. G. Relation of stuttering to mental fatigue. *J. exp. Psychol.*, 1934, **17**, 574–584.
4. Bills, A. G. Some causal factors in mental blocking. *J. exp. Psychol.*, 1935, **18**, 172–185.
5. Bills, A. G. Fatigue, oscillation and blocks. *J. exp. Psychol.*, 1935, **18**, 562–573.
6. Bills, A. G. A new psychergometer. *J. gen. Psychol.*, 1936, **14**, 487–489.
7. Bills, A. G., & Shapin, M. Mental fatigue under automatically controlled rates of work. *J. gen. Psychol.*, 1936, **15**, 335–347.
8. Bills, A. G. Blocking in mental fatigue and anoxemia compared. *J. exp. Psychol.*, 1937, **20**, 437–452.
9. Brown, W. Practice in associating color names with colors. *Psychol. Rev.*, 1915, **22**, 45–54.
10. Du Bois, P. The sex difference on the color-naming test. *Amer. J. Psychol.*, 1939, **52**, 380–383.
11. Fisher, R. A. *Statistical methods for research workers.* (7th ed.) London: Oliver and Boyd, 1938.
12. Garrett, H. E., & Lemmon, V. W. Analysis of several well-known tests. *J. appl. Psychol.*, 1924, **8**, 424–438.
13. Hansen, C. F. Serial action as a basic measure of motor capacity. *Psychol. Monogr.*, 1922, **31**, 320–382.
14. Hollingworth, H. L. Articulation and association. *J. educ. Psychol.*, 1915, **6**, 99–105.
15. Husband, R. W., & Ludden, M. J. Sex differences in motor skills. *J. exp. Psychol.*, 1931, **14**, 414–422.
16. Jersild, A. T. Mental set and shift. *Arch. Psychol.*, N. Y., 1926–27, **14**, 1–81.
17. Ligon, E. Genetic study of color naming and word reading. *Amer. J. Psychol.*, 1932, **44**, 103–122.
18. Lund, F. H. Role of practice in speed of association. *J. exp. Psychol.*, 1927, **10**, 424–434.
19. Martinson, B. Study of brain potentials during mental blocking. *J. exp. Psychol.*, 1939, **24**, 143–157.
20. Pearson, K. On the dissection of frequency curves. *Phil. Trans. Roy. Soc.*, 1894 (No. 1 of: Contributions to the mathematical theory of evolution).
21. Seashore, C. E. A method of measuring mental work: the psychergograph. *Univ. Ia. Stud. Psychol.*, 1902, **3**, 1–17.
22. Seashore, R. H. Techniques for measuring serial action. *J. exp. Psychol.*, 1928, **11**, 45–55.
23. Stroop, J. R. Studies of interference in serial verbal reactions. *J. exp. Psychol.*, 1935, **18**, 643–662.
24. Telford, C. W. Differences in response to colors and their names. *J. genet. Psychol.*, 1930, **37**, 151–159.
25. Vaughn, J., & Strobel, E. J. A colored light exposure apparatus. *J. gen. Psychol.*, 1940, **23**, 443–444.
26. Warren, N., & Clark, B. Blocking in mental and motor tasks during a 65-hour vigil. *J. exp. Psychol.*, 1937, **21**, 97–105.
27. Weaver, H., & Roberts, J. A discriminative serial action apparatus. *J. exp. Psychol.*, 1942, **31**, 171–176.
28. Woodworth, R. S., & Wells, F. L. Association tests. *Psychol. Monogr.*, 1911, **13**.

REPETITIVE AND ALTERNATIVE RESPONSES AND SEQUENCES OF ERRORS IN THE DISCRIMINATION OF COLOR MASS

BY B. R. PHILIP

Fordham University

Psychologists have frequently called attention to the complexity of the judgment process and have enumerated factors such as occur in space and time errors, contrast effects in judgment, and the series effect, to mention but a few, which exert small but significant influences on psychophysical thresholds.

This paper considers the influence of certain cues based upon the knowledge of previous judgments in discriminations near the threshold. It investigates tendencies after a judgment has been made to alternate the response when the subsequent stimulus is presented; or, on the other hand, to repeat the response previously made when the subsequent stimulus is presented. In addition, it investigates *S*'s tendency to make errors in sequence when disturbed by the possibility that an error has just been committed.

Material

The material of this experiment is similar to that which has been used in previous work of the author to study the discrimination of color mass (1, 2). It consists of a series of cards, on each of which is a number of dots printed in two colors, one of which predominates. The subject is told to judge the predominating color merely by the mass effect and without counting the number of dots. As the judgments may be made with precision and celerity, and the cards may be quickly exposed, easily changed in serial order, color pattern and complexity, this type of material is excellent for psychophysical work.

The color mass cards of this experiment had 105 dots arranged in 7 rows of 15 dots apiece, the dots being one-eighth inch in diameter, equally spaced at one-eighth inch intervals. There were only two colors, Blue and Purple, as previous work had shown that the difficulty of this color combination is satisfactory. There were four series of cards of varying difficulty, the number of dots of the predominating color being respectively 61, 59, 57, and 55; consequently the number of the dots of the other color were respectively 44, 46, 48, and 50. A series, say the 59 series, consisted of 100 cards, 50 of which had Blue as the predominating color, and 50 had Purple predominating. Each card in a series had 59 dots of the predominating color and 46 of the other. The dots were arranged in 10 different patterns; as they might be randomly presented upside down, in reality there were 20 different patterns of dots in the series. This variation in the dot pattern was required in order to prevent identification of the card from its pattern or dot arrangement, rather than from the color mass. In addition to these 20 different dot patterns, a further control to prevent memorization of the order of presentation of the predominating color (*i.e.*, memorization of the sequence of correct answers, as B, B, P, B, P, P, etc.) was obtained by using a lengthy sequence of 100 cards. Moreover, 5 completely different sequences were made up for each of the 4 sets. Separate packs of cards were carefully assembled according to these sequences, and the scoring charts were devised to show these sequences and the responses given.

Procedure

The cards were placed on a neutral gray mat, about 30 inches square, in front and on both sides of which similar mats were placed in an upright screening position. The front mat had an aperture five by two inches which permitted the easy insertion and withdrawal of the cards on which judgments were being made. All work was done by Macbeth Daylight illumination, under strong and even light. As fast as a judgment was made, the judged card was withdrawn and a new one inserted; meanwhile the response was being recorded on a special scoring sheet. *S*s worked quickly and readily became adapted to experimental conditions so that a series could be judged in from fifteen to twenty minutes. A short rest which averaged about four minutes was given between series, during which the new material was being prepared. Very few complaints were made by *S*s of fatigue or eye strain.

Instructions

The instructions were as follows:

"When a card is placed before you state whether Blue or Purple predominates merely by saying 'Blue,' or 'Purple.' Do not try to count the dots; make your judgments without undue hesitation. When in doubt, guess." *S*s were then shown the experimental routine and a few preliminary judgments were made. Acting alternately as *S* or *E*, 12 graduate students in psychology, 7 men and 5 women completed all the series. As a rule two *E*s were required for each *S*, one to present the cards and the other to score. Scoring was facilitated by the fact that each card had coded letters on the reverse side to indicate the type of card being presented, and the score sheets carried in correct order the same coding. All packs of cards were carefully checked beforehand with the scoring sheet. No difficulty was presented in scoring and the responses could be easily taken as fast as given.

The score sheets contained the letters B or P printed in proper sequence to indicate the presentation of a Blue or Purple stimulus card. If the response were correct a stroke was drawn through the letter; if incorrect, the letter was circled. On the scoring sheet there were 10 rows of 10 letters each, opposite each row the correct and incorrect responses were separately totalled for both B and P.

The sequences in which the stimuli were given are of importance for this study. The stimuli were given in a random order subject to the following limitations: First, in every group of 20 stimuli there was an equal number of B's and P's. Secondly, there were no more than 4 consecutive stimuli of the same type. Hence the stimuli could be alternated in the sequence BPBP, or PBPBPB, etc. or they could be repeated in doublets, triplets, and quadruplets, *e.g.*, BB, PPP or BBBB. Table 1 shows how these repetitive stimuli were distributed fairly equally for each of

TABLE 1

Frequency of Repetition Types in a Series

Type	BBP	PPB	BBBP	PPPB	BBBBP	PPPPB
	56	44	13	15	2	5
	100		28		7	

TABLE 2

Frequency of Alternative Types in a Series

	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
BP	8	4				1		2		1				16
PB	1	9	7		3		1	3	2	2			1	29
Total	9	13	7		3	1	1	5	2	3			1	45
Cumulative Frequency	45	36	23	16	16	13	12	11	6	4	1	1	1	

the 5 sequences comprising a series (500 stimuli per series). Thus in each series there were 100 doublets, 28 triplets, and 7 quadruplets.

The alternation types consist of from 2 to 14 members, the distributions of which are given in Table 2. The cumulative frequency is also given to show the relative frequency with which alternations should be expected in a series at every point starting from the first stimulus of the alternation.

Results

Relative Difficulty of the Series

At the completion of the test the cards were checked for totals of Rights and Wrongs for each color. The percentage of accuracy as determined by the formula $100 \cdot \frac{R - W}{R + W}$ was computed for each color, from which the average percentage score was obtained. Each sheet yielded 100 judgments, and there were 5 sequences for each of the 4 series, hence there were 24,000 judgments in all from the 12 *Ss*. The percent accuracy score was averaged for each series. The results are given in Table 3. It is to be noted that *Ss* differed

TABLE 3

Percentage of Accuracy for the Different Series

Subject		Series 61	Series 59	Series 57	Series 55	Average	Blue	Purple
1		96.4	88.0	58.4	81.6	81.1	108	81
	P	.87	.99	*x*	.01			
2		95.6	85.6	81.6	3.2	65.5	180	165
	P	.80	.79	.01	*x*			
3		13.6	−4.4	75.6	28.8	28.4	325	391
	P	*x*	*x*	.01	*x*			
4		74.0	68.8	85.2	19.2	61.8	199	183
	P	.17	.47	.03	*x*			
5		94.8	68.8	88.4	56.4	77.1	145	84
	P	.56	.02	<.01	*x*			
6		58.8	49.2	82.0	26.4	54.1	188	271
	P	.01	*x*	<.01	*x*			
7		59.6	45.6	74.8	39.6	54.9	217	234
	P	<.01	*x*	.04	*x*			
8		82.4	62.8	94.4	18.0	64.4	195	161
	P	.86	.12	.83	*x*			
9		82.4	55.6	93.6	48.0	69.9	151	150
	P	.01	*x*	.90	*x*			
10		87.6	44.4	82.0	51.6	67.4	163	163
	P	.91	*x*	.21	*x*			
11		76.4	81.6	96.0	16.4	67.5	133	192
	P	.03	.73	.95	*x*			
12		80.8	72.0	54.8	6.8	53.6	303	161
	P	<.01	.15	*x*	*x*			
Average		75.2	59.8	80.6	33.0	62.1	192.25	186.33

Note: Under the columns headed Blue and Purple is listed the number of errors made on 1000 stimuli of the type given. Under the percentage of accuracy for each series are given the *P* values corresponding to the χ^2 test for agreement with the Poisson series; .01 implies that the *P* value is between .005 and .01; *x*, that it is less than .005.

considerably in performance on this test, that the Purple judgments were not significantly easier than the Blue judgments (18.63 percent error for P as compared with 19.22 percent error for B: critical ratio of difference = 0.35) and that the series in increasing order of difficulty were as follows, 57, 61, 59, and 55, the corresponding percentage accuracy being 80.6, 75.2, 59.8, and 33.0. This divergence from the difficulty that one might expect from the physical arrangement of dots was due to the fact that the 57 series had not been

TABLE 4

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN PERCENTAGE ERRORS AT VARIOUS POSITIONS

	Repetitive Type						
Positions	2–1	3–2	4–3	5–4	3–1	4–1	5–1
BBP	4.29	−5.44			−1.12		
PPB	3.45	−3.36			0.09		
Total—Doublets	5.39	−6.43			−0.77		
BBBP	4.25	3.56	−6.50		7.82	1.32	
PPPB	3.53	2.62	−5.33		6.11	0.91	
Total—Triplets	5.55	4.37	−8.42		9.94	1.54	
BBBBP	0.41	2.61	−1.31	−3.02	3.03	1.70	−1.37
PPPPB	0.23	0.11	2.10	−5.06	0.35	2.45	−2.62
Total—Quadruplets	0.42	1.55	1.01	−5.96	1.97	2.98	−3.00

	Alternative Type					
Positions	2–1	3–2	5–4	7–6	8–6	9–8
BP	2.65	−3.42	1.43	1.74	2.85	−1.25
PB	4.65	−3.68	0.17	0.79	−0.09	−1.09
Total	4.87	−5.04	1.08	1.08	2.45	−1.69

A minus sign before the critical ratio implies that the percentage of error in the second position is less than that in the first position.

printed as well as the other three. On three of the series the Blues and the Purples had been fairly well equated for intensity, but on the 57 series the Purples differed somewhat in intensity from the Blues, which characteristic *Ss* used as a cue. Since, in material of this type there is a logarithmic relationship between percentage of predominating color and percentage of accuracy (*cf.* 1), if the series 57 fitted in with the others it should have yielded approximately 45 percent accuracy rather than the 80.6 percent obtained. While this misprinting of the material rendered the 57 series useless for the purpose of determining psychophysical limens, it did not affect the value of this series in the determination of the main factors which were investigated.

The Repetitive Type of Response

The primary interest of this study was to compare the tendency to alternate a previous judgment with the tendency to repeat that judgment. Thus if the response B were given at any point in the test, an alternative response for the next stimulus would be P, and a repetitive response would be B.

The repetitive type of response was first investigated. A tabulation was made for both color stimuli of the errors at each position

TABLE 5

RELATIVE FREQUENCY (p) OF ERROR ACCORDING TO POSITION IN REPETITION TYPE

Position		I	II	III	IV	V	Total Errors
BBP	p	0.157	0.202	0.146			2688
	σ	0.007	0.008	0.007			
PPB	p	0.162	0.203	0.163			2112
	σ	0.008	0.009	0.008			
Total	p	0.159	0.203	0.153			4800
	σ	0.006	0.006	0.005			
BBBP	p	0.141	0.234	0.324	0.168		624
	σ	0.014	0.017	0.019	0.015		
PPPB	p	0.129	0.200	0.258	0.146		720
	σ	0.013	0.015	0.016	0.013		
Total	p	0.135	0.216	0.289	0.156		1344
	σ	0.009	0.011	0.012	0.010		
BBBBP	p	0.135	0.156	0.313	0.229	0.073	96
	σ	0.035	0.037	0.047	0.043	0.029	
PPPPB	p	0.167	0.175	0.179	0.258	0.088	240
	σ	0.024	0.024	0.025	0.028	0.018	
Total	p	0.158	0.170	0.217	0.250	0.083	336
	σ	0.020	0.020	0.022	0.024	0.015	

of the doublets, triplets, and quadruplets, and of the first position immediately following these types. Thus, for a triplet in which Blue was the stimulus, the correct responses for each position of the triplet and the position immediately following would be BBBP. The giving of the contrary response at any position would constitute an error for that position. These tabulations were made for all different repetitive types and for all four series. The results are presented in Tables 5 and 6 in percentile form, and are indicated in graphical form in Fig. 1. It is obvious that the percentage of error rises continuously with position throughout the doublet, triplet, and quadruplet, and immediately drops when the stimulus is changed.

TABLE 6

RELATIVE FREQUENCY OF ERROR ACCORDING TO POSITION FOR REPETITION TYPES FOR DIFFERENT SERIES

		I	II	III	IV	V	Total
Doublet Series	61	.105	.130	.089			1200
	59	.149	.206	.179			1200
	57	.070	.131	.050			1200
	55	.314	.344	.306			1200
Triplet Series	61	.074	.116	.244	.107		336
	59	.146	.202	.315	.143		336
	57	.059	.149	.181	.057		336
	55	.262	.396	.414	.318		336
Quadruplet Series	61	.083	.060	.119	.143	.024	84
	59	.202	.202	.190	.298	.048	84
	57	.036	.048	.143	.214	.036	84
	55	.310	.369	.417	.346	.226	84

Thus in the type BBBP the errors increase with position when the response B is called for, and decrease abruptly when response P is required. The progressive rise in error with position occurs for

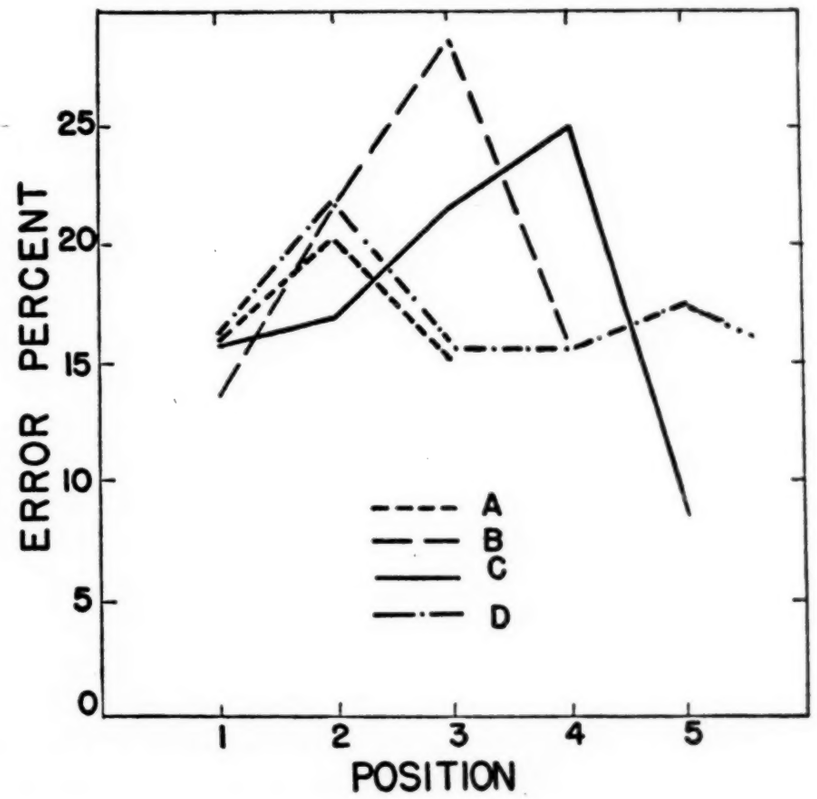


FIG. 1. Percentage of errors according to position for repetition and alternation types of response. Repetition types: doublet, *A*; triplet, *B*; quadruplet, *C*. Alternation type, *D*. Only the first five positions of this type are shown. To be noted in *A*, *B*, and *C*, is the progressive increase of percent error with position as long as the stimuli are repeated; and the decrease when the stimulus is changed.

practically all types, stimuli and series, the quadruplet forms showing some slight divergences, which might be expected as the stimuli in that case were relatively few. Obviously some tendency is operating continuously and increasingly throughout the three repetitive types to cause an erroneous response.

That this increase in error is significant is shown in Table 4, which gives the critical ratios as determined by the formula $\frac{\text{Diff.}}{\sigma \text{ diff.}}$. The standard errors of the percentages calculated by the formula $\sqrt{\frac{pq}{N}}$, where p = percentage of error, $q = 1 - p$, and N = the number of judgments, are given in Table 5. It is readily seen that for the doublets and the triplets, significantly lesser percentages of error are made when the stimuli are changed at the end of the doublet or triplet. The quadruplet types yield increasing amounts of error percentage, but the differences from position to position are significant only when the stimuli are repeated through several positions (*e.g.*, positions 4–1). Significant differences are obtained, however, when the stimuli are changed (*e.g.*, positions 5–4). In view of the relatively small number of judgments of the quadruplet type this deviation from statistical significance is perhaps to be expected. On the other hand the regularity with which the phenomenon occurs even when the data are fractionated for series, Table 6, and for sequence type, Table 5, gives some confidence in their significance.

It must be remembered that these judgments are difficult to make, and for the most part are in the neighborhood of the discriminatory threshold. Even when correctly made they are seldom accompanied by any great degree of assurance. To aid him in his judgment *S* is eager to grasp additional cues, the influence of which is out of proportion to the precision with which they may be perceived. Furthermore a form of the stimulus error is present. Rather than base his judgments on completely objective criteria, *S* tends to discriminate the stimuli, in part at least, by what he imagines they might be, even before they are presented. In other terms he tends to anticipate the stimuli on the basis of some kind of probability. In the course of the experimental session, *S* becomes indirectly aware of the relative frequency of occurrence of doublets, triplets, quadruplets, and of the alternative forms. He does not estimate this frequency in numbers, and if asked to do so, would probably be considerably in error, nevertheless his judgments are subtly influenced by this vague statistical knowledge. Of importance too is the fact that his subsequent judgments are influenced by his prior judgments. In other terms he feels that having made a

judgment on a stimulus of a certain type, the probability of occurrence of these stimuli will thereby be lessened; in reality, of course the selection of the stimulus that will be presented is in large part independent of previous selections being determined only by the conditions outlined above. Goodfellow has pertinently discussed this point (**3**).

As a consequence of his lack of confidence in these difficult judgments and of his tendency to use some type of probability as an aid in framing his judgments, *S* is prone to render a judgment counter to his discriminatory ability—his subjective attitude is antagonistic to his objective judgment. To illustrate, in a triplet of the form PPPB, *S* judges the first position as 'P' fairly accurately and confidently. His ability does not change with the second position, but, owing to his anticipation of a change in stimulus, which does not occur, his confidence in the correction of his judgment is diminished, and he is prone to respond 'B' to some of the stimuli, thus increasing his percentage of error in the second position. In the third position, if he has made two successive 'P' judgments, he feels that the stimuli must now be changed, and he is far more prone to respond with 'B,' thus further increasing his percentage error. In the fourth position this tendency to look for a change in stimulus aids him in framing the judgment and hence his percentage of error diminishes. In short, while *S*'s real discriminatory ability does not change with position, his tendency to alter his judgments in repetitive types of stimuli does increase with position, to the detriment of his accuracy.

The Alternative Type of Response

In the study of the alternative types a tabulation was also made of the errors at each position of the alternation, and also of the first position immediately following, *i.e.*, of the first response of the doublet, triplet, or quadruplet which follows the alternation. Separate tabulations were made for each series and for each color sequence, since the alternation could be of the form BPBP or of the form PBPB, etc. The results of the series tabulations are not presented here; the tabulations for the sequence types are presented in Table 7 as percentages of the error in judging the stimuli, and are shown in Figs. 1 and 2. Obviously the results are not so clear-cut in the case of the alternations as they are for the repetitions. What is evident is the presence of maxima at positions 2, 5, and 8, and a terminal maximum. However, as the cumulative frequency tabulation starting from the past position, given in Table 2, shows, stimuli occur relatively infrequently beyond position 9, hence the form of the curve beyond that position is probably not significant.

TABLE 7

RELATIVE FREQUENCY (p) OF ERROR ACCORDING TO POSITION IN ALTERNATIVE TYPE

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
PB	p	0.166	0.205	0.155	0.160	0.189	0.149	0.190	0.220	0.185	0.167	0.118	0.104	0.188	0.229
	σ	0.0100	0.0108	0.0099	0.0121	0.0163	0.0148	0.0189	0.0199	0.0198	0.0241	0.0269	0.0441	0.0565	0.0606
BP	p	0.151	0.245	0.156	0.141	0.135	0.198	0.167	0.194	0.146	0.167	0.125			
	σ	0.0129	0.0155	0.0185	0.0251	0.0246	0.0287	0.0269	0.0329	0.0294	0.0538	0.0477			
Total	p	0.161	0.219	0.156	0.157	0.176	0.161	0.183	0.214	0.174	0.167	0.120	0.104	0.188	0.229
	σ	0.0079	0.0089	0.0087	0.0109	0.0137	0.0133	0.0155	0.0170	0.0165	0.0220	0.0235	0.0441	0.0565	0.0606

The most interesting fact is that the first three positions of the alternation curve very closely resemble the form of the doublet curve. This difference in error percentage between the first and the second positions and second and third positions is statistically reliable, the C.R.'s being respectively 4.87 and − 5.04. *Cf.* Table 4. The error percentage difference between the minimum at the 6th position and the maximum at the 8th position approaches significance (C.R. = 2.45).

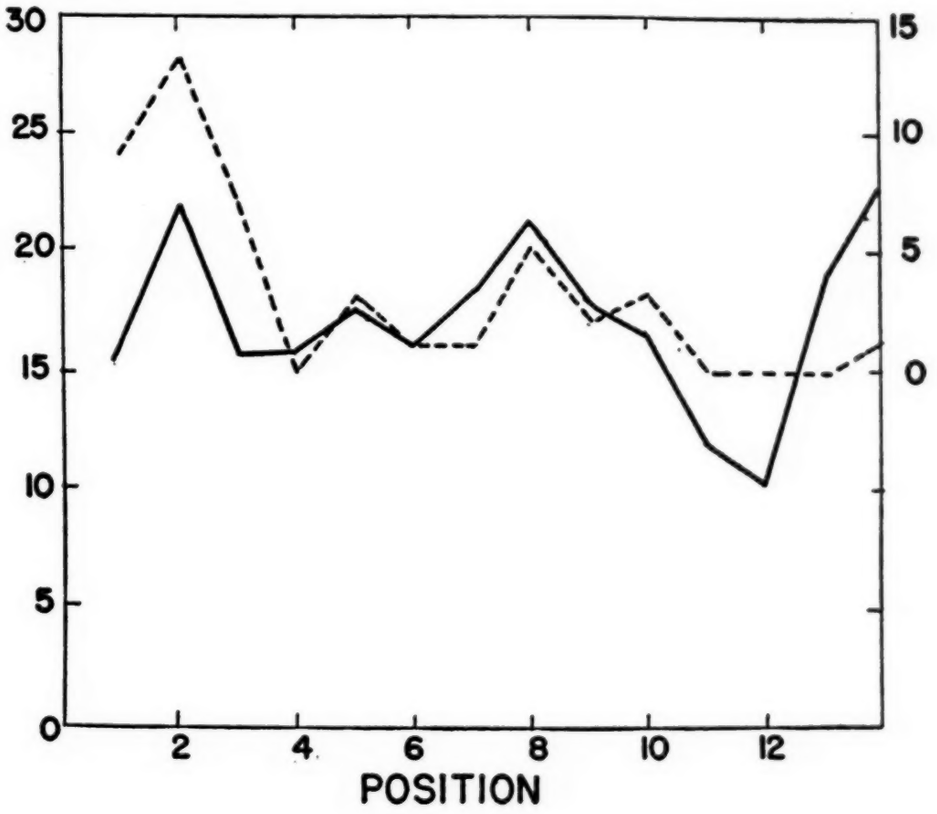


FIG. 2. Comparison of the percent error in the alternative types (continuous line) with the number of alternative sequences terminating at the positions indicated (broken line). Abscissæ indicate position; ordinates, percent error, for the continuous line on the left; and for the broken line, on the right, number of sequences.

In the repetitive type of stimuli we found a clear-cut relationship between position and percentage of error, the error percentages increasing steadily with position. Hence in the alternative type of stimuli one might be prone to expect maxima and minima to alternate with position. Obviously this is not so; even where an alternation does occur (in the first three positions), it is rather contrary to what one might assume from the findings of the repetitive type of stimuli; there errors increase with position because *S* tends to *change* his judgment. The rise in error percentage at the second position and

the drop in the third position in the alternation type implies that *S* tends to *repeat* his judgment, at least for the first three positions.

Furthermore, if on the curve of error percentage with position is also plotted for sake of comparison a curve showing the number of alternations which terminate at the various positions, it can readily be noted that the general trend of both curves is similar (Fig. 2).

This fact probably implies that in the case of the alternative as in the case of the repetitive type of stimuli *S* becomes vaguely aware of the experimental frequency of the sequence types, and sometimes attempts to adjust his judgments to this knowledge, rather than to discriminate solely on the basis of the color mass of the stimuli. In other terms, the indirect cues that he obtains from his experience with the experimental set-up detrimentally affect his judgment, impelling him to go counter to his judgment based on discriminatory ability. This impulse is the more urgent with a consequent increase in error the longer a series continues. At times indeed this impulse may bring greater accuracy, whenever his tendency to anticipate the sequence of the stimuli coincides with their actual sequence; *e.g.*, when the stimuli are changed after a repetitive type.

What is apparent then, is not so much that there is a conflict between an alternative tendency and a repetitive tendency in making a judgment, as an attempt to utilize a vague statistical knowledge of the sequence arrangement as a cue in framing the judgments. Cues thus obtained are often at variance with the judgments based on discriminatory ability, and progressive increases of error percentage with position arise from resultant tendencies to alternate when repetitions are required and to repeat when alternations are required.

Consequently one could predict the relative strength of the alternative and repetitive tendencies, based upon the errors made by both, by determining the relative frequency with which alternations and repetitions are called for by the experimental situation. No precise figure can be obtained, but an approximation may be ascertained by noting the relative frequencies of alternations and repetitions in the second position of a sequence. It will be recalled that in each of the four series there were 100 doublets, 28 triplets and 7 quadruplets, making 135 repetitions in the second position. Correspondingly there were only 45 alternations. However the ratio 135 : 45 does not give a correct value for the relation between repetitions and alternations, owing to the varying lengths of the sequences. When the sequences are weighted for length, the ratio obtained is 312 : 230, *i.e.*, 1.36.

In an attempt to determine the actual number of errors of the repetitive type and those of the alternative type, Table 8 is presented.

TABLE 8

FREQUENCY OF ERRONEOUS REPETITIONS AND ALTERNATIONS

	Series 61			Series 59			Series 57			Series 55		
	a	*b*	*a/b*	*a*	*b*	*a/b*	*a*	*b*	*a/b*	*a*	*b*	*a/b*
Subject 1	14	3	4.67	40	13	3.08	105	50	2.10	45	22	2.05
2	12	3	4.00	39	19	2.05	35	28	1.25	188	128	1.47
3	166	117	1.42	203	138	1.47	57	27	2.11	143	95	1.51
4	58	37	1.57	74	36	2.06	20	25	0.80	148	118	1.26
5	13	8	1.63	71	39	1.82	24	15	1.67	87	58	1.50
6	96	58	1.66	120	69	1.74	47	32	1.47	122	102	1.20
7	99	43	2.30	132	63	2.10	72	37	1.94	123	83	1.48
8	45	28	1.61	101	41	2.46	15	8	1.87	171	105	1.63
9	57	18	3.16	103	53	1.94	16	7	2.28	117	69	1.70
10	17	6	2.83	118	67	1.76	47	24	1.96	83	64	1.30
11	56	28	2.00	58	21	2.76	9	7	1.28	153	106	1.44
12	36	27	1.33	61	24	2.54	90	64	1.41	196	116	1.69
Total	669	376		1120	583		537	324		1576	1066	
Ratio	1.78			1.92			1.66			1.48		

a. Erroneous repetitions.
b. Erroneous alternations.

It was obtained as follows: From the score sheets for each *S* and for each series a computation was made of:

(*a*) The number of alternations which should have been made and were not made; hence an erroneous repetition of the previous response was given.

(*b*) The number of incorrect alternations, *i.e.*, when *S* alternated instead of continuing the response as he should have done.

The ratios, 48 in all, between (*a*) and (*b*) were worked out for each *S* and for each series. Only one of these ratios is less than 1.0. The average ratio of the group, found by dividing the errors of the repetitive type and those of the alternative type, is 1.56. The agreement with the ratio 1.36 given above, which shows approximately the relation between the stimuli that were repeated and those that were alternated, seems satisfactory.

Sequences of Errors

In the course of the experimentation it was noticed that *Ss* sometimes became emotionally disturbed when making their judgments. These disturbances, chiefly noted in the harder series, seemed to cause sequences of errors. In an attempt to analyze the tendency to repeat errors a tabulation was compiled of the frequencies of errors, one at a time, two at a time, etc., made by every *S* in each series. In Table 9 the total tabulations for the series are given.

TABLE 9

COMPARISON OF FREQUENCY OF ERRORS WITH POISSON SERIES

	Series 61		Series 59		Series 57		Series 55	
	Errors	Poisson	Errors	Poisson	Errors	Poisson	Errors	Poisson
0	5264	5120.6	4796	4590.9	5419	5327.6	4009	3611.2
1	427	650.3	674	973.3	376	527.4	794	1365.0
2	83	41.6	128	103.2	65	26.4	306	258.0
3	25	1.8	52	7.3	15	0.9	101	32.5
4	10	0.1	15	0.4	5		36	3.1
5	4		5		2		13	0.2
6	—		4				5	
7	—		—				5	
8	1		—				1	
9			1					
Mean = m	0.127		0.212		0.099		0.378	
e^{-m}	0.88073		0.80896		0.96574		0.68523	
χ^2	229.9		101.3		16.7		728.7	

The Poisson series seems adequately suited to describe the fluctuations of samplings of errors. In general if a number can take the values 0, 1, 2, 3 $\cdots$ k, and the relative frequencies with which the values occur are given by the series

$$e^{-m}\left(1,\ m_1 \frac{m^2}{!2} \frac{m^3}{!3} \cdots \frac{m^k}{!k} \cdots\right),$$

where m is the mean and $!k$ stands for factorial $k = k(k-1)(k-2) \cdots 1$, then the number is distributed in the Poisson series. This series is used in considering events that rarely happen, and the successive terms give the frequency of occurrences of 0, 1, 2, 3, etc., errors on the basis that each error be not prejudiced by what has previously occurred (4). As regards the data of this experiment it seems plausible that, should the error distribution experimentally found significantly differ from the values calculated for the Poisson series, one might assume that 'previous occurrences' had affected the probability of error. Among the 'previous occurrences' was undoubtedly the committing of errors to the detriment of S's subsequent judgments. His knowledge of these errors was of course uncertain, but this very uncertainty increased his disturbance.

The χ^2 test was used to determine the conformity of the experimentally found values with the calculated values. The corresponding P values are given in Table 3. It can readily be seen that the conformity of the Poisson series is best in the easier series and that there is no significant agreement in the 55-series, except for the

best subject, who found this series relatively easy. A scatter-diagram of the variables percentage of accuracy and P values is definitely curvilinear. It may be noted that when the discrimination is below the threshold (percentage of accuracy is less than 50), the Poisson series does not fit the curve showing the error distribution. That this divergence is due chiefly to the occurrence of errors in

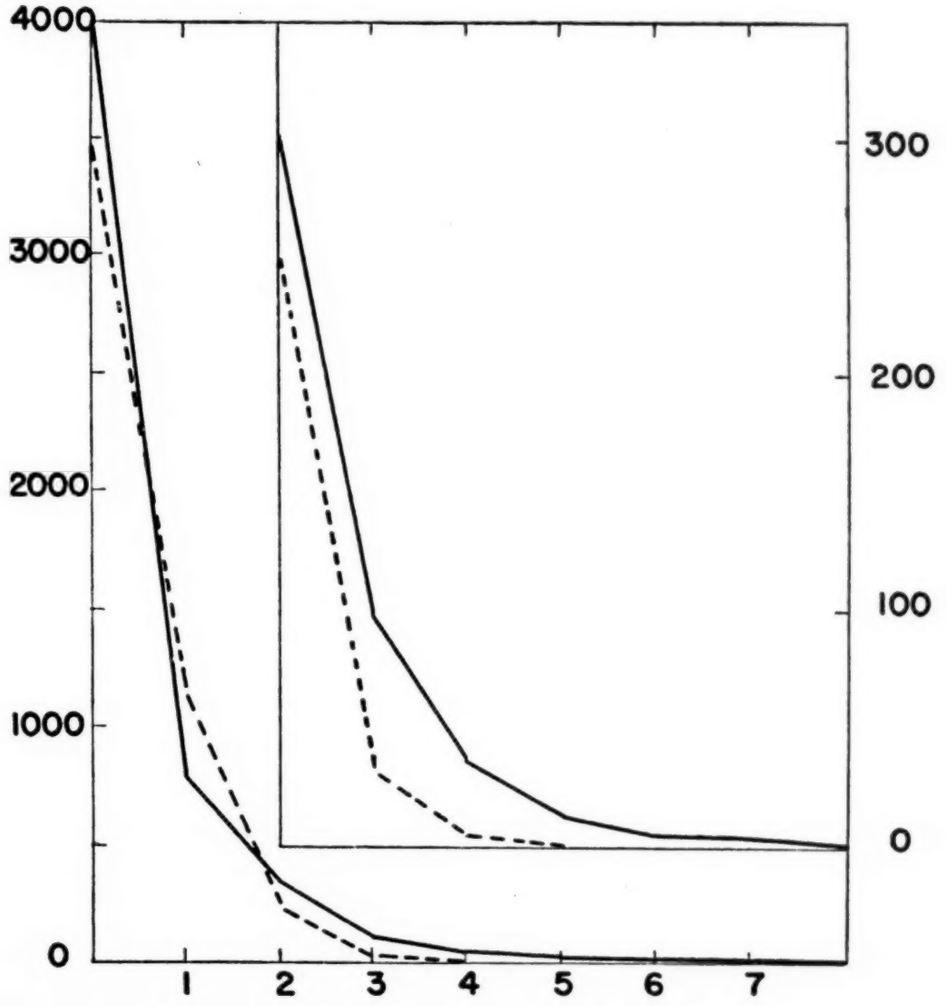


FIG. 3. Total frequency of sequences of errors, Series 55, fitted to a Poisson series. The broken line gives the theoretical frequencies of the Poisson series, and the continuous line, the frequency of occurrences of sequences of errors. Abscissæ indicate errors in a sequence; ordinates, the frequency of occurrence, on the left for the complete curve, on the right for the enlarged section of the same curve.

sequence may be seen from Fig. 3 which compares for series 55 the frequency of errors and the corresponding Poisson values. Table 9 presents these data for all four series. It may be noted individual errors occur less frequently and that sequences of two, three, four, etc., errors occur much more frequently than chance expectancy, as given by the Poisson series. The probable explanation for this divergence, in part at least, is that S becomes confused when making

TABLE 9

COMPARISON OF FREQUENCY OF ERRORS WITH POISSON SERIES

	Series 61		Series 59		Series 57		Series 55	
	Errors	Poisson	Errors	Poisson	Errors	Poisson	Errors	Poisson
0	5264	5120.6	4796	4590.9	5419	5327.6	4009	3611.2
1	427	650.3	674	973.3	376	527.4	794	1365.0
2	83	41.6	128	103.2	65	26.4	306	258.0
3	25	1.8	52	7.3	15	0.9	101	32.5
4	10	0.1	15	0.4	5		36	3.1
5	4		5		2		13	0.2
6	—		4				5	
7	—		—				5	
8	1		—				1	
9			1					
Mean = m	0.127		0.212		0.099		0.378	
e^{-m}	0.88073		0.80896		0.96574		0.68523	
χ^2	229.9		101.3		16.7		728.7	

The Poisson series seems adequately suited to describe the fluctuations of samplings of errors. In general if a number can take the values 0, 1, 2, 3 $\cdots$ k, and the relative frequencies with which the values occur are given by the series

$$e^{-m}\left(1,\ m_1 \frac{m^2}{!2} \frac{m^3}{!3} \cdots \frac{m^k}{!k} \cdots\right),$$

where m is the mean and $!k$ stands for factorial $k = k(k-1)(k-2) \cdots 1$, then the number is distributed in the Poisson series. This series is used in considering events that rarely happen, and the successive terms give the frequency of occurrences of 0, 1, 2, 3, etc., errors on the basis that each error be not prejudiced by what has previously occurred (**4**). As regards the data of this experiment it seems plausible that, should the error distribution experimentally found significantly differ from the values calculated for the Poisson series, one might assume that 'previous occurrences' had affected the probability of error. Among the 'previous occurrences' was undoubtedly the committing of errors to the detriment of *S*'s subsequent judgments. His knowledge of these errors was of course uncertain, but this very uncertainty increased his disturbance.

The χ^2 test was used to determine the conformity of the experimentally found values with the calculated values. The corresponding P values are given in Table 3. It can readily be seen that the conformity of the Poisson series is best in the easier series and that there is no significant agreement in the 55-series, except for the

best subject, who found this series relatively easy. A scatter-diagram of the variables percentage of accuracy and P values is definitely curvilinear. It may be noted that when the discrimination is below the threshold (percentage of accuracy is less than 50), the Poisson series does not fit the curve showing the error distribution. That this divergence is due chiefly to the occurrence of errors in

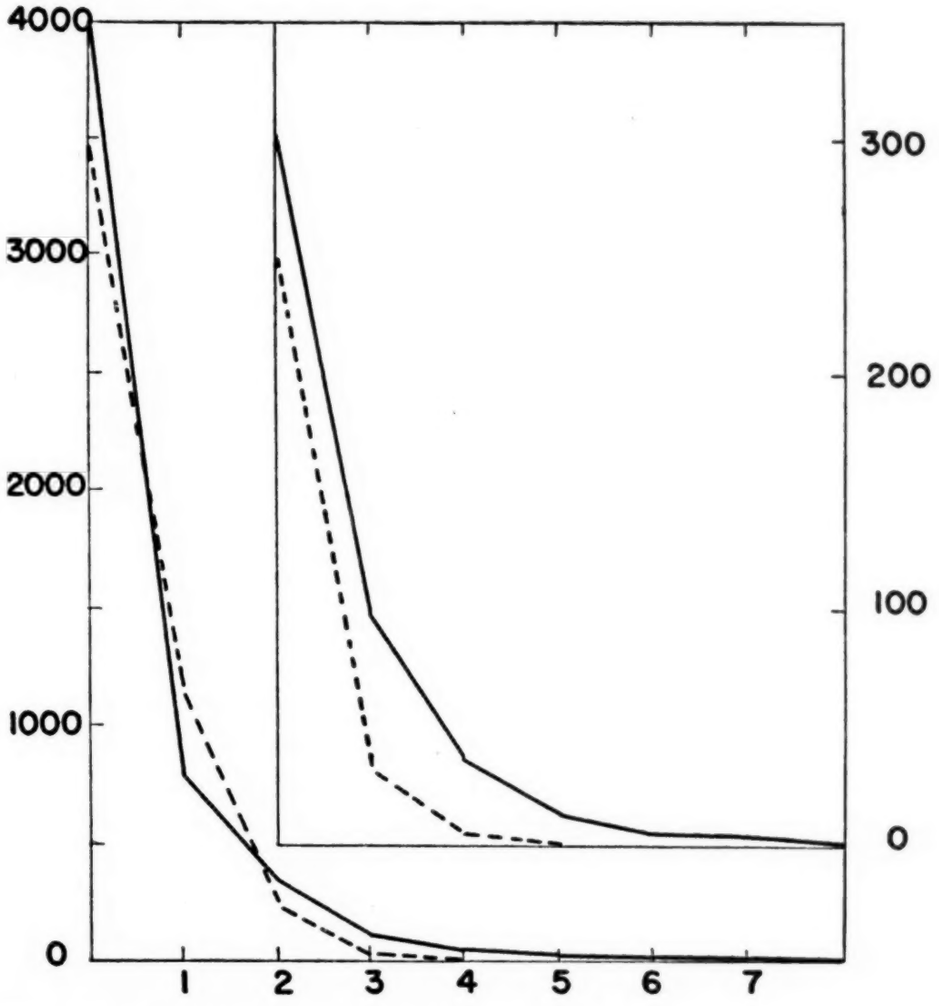


FIG. 3. Total frequency of sequences of errors, Series 55, fitted to a Poisson series. The broken line gives the theoretical frequencies of the Poisson series, and the continuous line, the frequency of occurrences of sequences of errors. Abscissæ indicate errors in a sequence; ordinates, the frequency of occurrence, on the left for the complete curve, on the right for the enlarged section of the same curve.

sequence may be seen from Fig. 3 which compares for series 55 the frequency of errors and the corresponding Poisson values. Table 9 presents these data for all four series. It may be noted individual errors occur less frequently and that sequences of two, three, four, etc., errors occur much more frequently than chance expectancy, as given by the Poisson series. The probable explanation for this divergence, in part at least, is that S becomes confused when making

difficult judgments, loses confidence in his discriminatory ability especially when he has made an error, and hence tends to make two or more errors in sequence.

Introspections

The findings here presented were substantiated by the introspections of *S*s, and their responses to questions when the experimental sessions were all completed. Nearly all spoke of the difficulty of the series and admitted lack of confidence in their judgments especially when they became confused. Of course they did not precisely know the relative frequency of the repetitive and alternative types, and on this point they were quite emphatic, but in a general way they learned the approximate set-up. Not one subject stated that this knowledge affected his judgment but some acknowledged that it may have done so. It is obvious that any utilization of the experimental arrangement of the stimuli as a cue in discrimination could only have been effected subliminally and without conscious effort.

Discussion

The results of this investigation seem to be in line with several others which have uncovered kindred effects of the influence of an unformulated knowledge of the experimental set-up as a cue in determining the response. Stated in other terms this investigation as well as those which are concerned with similar phenomena, is an attempt to study the adjustment of the response to the supposed probability of occurrence of the stimuli. On this score Brunswik aptly points out, in discussing probability as a determiner of rat behavior (**5**):

> In the natural environment of a living being, cues, means or pathways to a goal are usually neither absolutely reliable nor absolutely wrong. In most cases there is, objectively speaking, no perfect certainty that this or that will, or will not, lead to a certain end, but only a higher or lesser degree of probability. . . . The need to study ambiguity experimentally became especially urgent in the psychology of perception. The natural cues used by the perceptual system are, without exception, ambiguous in character. This in its turn, leads to ambiguity of the response as well. As indicated in recent studies, the perceptual system under certain conditions will learn to make use of new kinds of cues even if they do not prove to be highly reliable. . . .

Experimenters like Tolman and Brunswik (**6**) and Krechevsky (**7**) in the field of animal learning, and Humphreys in conditioning, have based their theories on the determining influences of probability factors. Thus Humphreys (**8**) states:

> A satisfactory alternative theory (of conditioning) requires a new interpretation of the roles of reinforcement and of non-reinforcement. One such alternative is that conditioned responses are a consequence of anticipated reinforcement, extinction of anticipated non-reinforcement, and

that the role of frequency in the repetition of reinforcement or non-reinforcement is by way of its influence on the subject's *expectations of the stimuli which are to appear.* (Italics mine.)

In a quite different field, that of psychophysics, Goodfellow (3) has called attention to the following caution:

The interdependence of an observer's judgments is due partly to his tendency to change his response because of his previous response. The basis of this tendency is probably the observer's vague feeling (due to a misconception of probability) that a chance event having already occurred is less likely to occur on the second trial.

Distinction should be made between situations in which a discrimination can be made, and where there is no guessing, and situations which evoke pure guesses, where the stimuli are presented only after the corresponding response has been made, or there is no objective means of making a correct judgment. The initial runs of mazes, and random matching experiments are illustrations of the latter type; ordinary psychophysical experiments, of the former. Even in psychophysical discrimination the probability factor, however, may come into play sooner or later, after *S* acquires some knowledge (not necessarily formulated) of the experimental set-up, which is used in an attempt to anticipate the stimuli. The alternative and repetitive tendencies which we have described, the expectations of the stimuli in conditioning, the patterns of response in psychophysics and parapsychology described by Goodfellow, and the ambiguous cues of Brunswik, alike fit to a greater or less degree in this category. In all of them the cues for response are antecedent to the stimuli about to be presented, but are derived from and are hence subsequent to, other stimuli and responses.

The tendency that we have noted in *S* to make errors in sequence when confused and disturbed, on the other hand, is due to his knowledge of responses already made, without significant reference to stimuli about to be presented. In this regard it resembles the contrast effect described by Fernberger (9), Arons and Irwin (10) and Preston (11), the series effect, and those effects which cause space and time errors. There is this difference: the effects just enumerated are due in part to an absorption of the immediate stimuli into a frame of reference which alters their stimulus value. The tendency to make errors in sequence is the result of a minor emotional disturbance, whose effect on the stimulus is to increase its difficulty, by rendering less precise its corresponding frames of reference.

(Manuscript received February 9, 1942)

BIBLIOGRAPHY

1. PHILIP, B. R. The discrimination of color mass. *Amer. J. Psychol.*, 1938, **51**, 127–136.
2. PHILIP, B. R. Intra-serial effects in the discrimination of color mass by the ranking method. *J. exp. Psychol.*, 1940, **27**, 125–135.

3. Goodfellow, L. D. The human element in probability. *J. gen. Psychol.*, 1940, **23**, 201–205.
4. Soper, H. E. Tables of Poisson's Exponential binomial limit. *Biometrika*, 1914, **10**, 25–35.
5. Brunswik, E. Probability as a determiner of rat behavior. *J. exp. Psychol.*, 1939, **25**, 175–198.
6. Tolman, E. C., & Brunswik, E. The organism and the causal texture of the environment. *Psychol. Rev.*, 1935, **42**, 43–77.
7. Krechevsky, I. 'Hypotheses' versus 'chance' in the pre-solution period in sensory discrimination. *Univ. Calif. Publ. Psychol.*, 1932, **6**, 27–44.
8. Humphreys, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eye-lid reactions. *J. exp. Psychol.*, 1939, **25**, 141–159.
9. Fernberger, S. W. Interdependence of judgments within the series for the method of constant stimuli. *J. exp. Psychol.*, 1920, **3**, 126–150.
10. Arons, L., & Irwin, F. W. Equal weights and psychophysical judgment. *J. exp. Psychol.*, 1932, **15**, 733–751.
11. Preston, M. G. Contrast effects and psychophysical judgments. *Amer. J. Psychol.*, 1936, **48**, 389–402.

THE DISSOCIATION OF BLOOD PRESSURE CONDITIONED RESPONSES UNDER ERYTHROIDINE

BY EDWARD GIRDEN

Brooklyn College *

Introduction

It has been demonstrated that learning established in curarized animals is repressed upon recovery from the drug but reappears spontaneously upon recurarization (**8**). The conditioned response (*CR*) to bell-stimulus consisted solely in the contraction of the isolated semitendinosus muscle. The dissociative effect produced by curare has been recently duplicated with erythroidine. More important, it has been demonstrated that the *CR* to light or bell in the *mild-drug-state* consisted of generalized 'emotional' activity of the entire organism (**6**).

One observing this phenomenon for the first time would be hard-put to distinguish it from the excited behavior evoked in the completely decorticate preparation or in the first stages of motor conditioning in normal animals.[1] The marked resemblance to the rage reaction suggested at once the presence of concomitant autonomic responses.

This paper is a report of the first of two investigations of the autonomic responses in the drug-state and, specifically, is concerned with the development of blood pressure (*BP*) *CR* during the erythroidine-induced dissociation. It should be kept clearly in mind that, so long as the *CR*-training was completely confined to the drug-state in normal animals (brain intact), none of these phenomena (*CR* of isolated muscle or generalized struggle-pattern) appeared after recovery from the drug. The repression was complete and the *CR* was evoked once again by the conditional stimulus *only* when the animal was reintroduced into the drug-state.

* Communication No. 45 from the Physiological Psychology Laboratory, University of Rochester, maintained by aid of the Research Council, American Otological Society. This research was performed during the writer's tenure as a Fellow of the John Simon Guggenheim Memorial Foundation, 1941–42.

Special aid from the Penrose Fund of the American Philosophical Society and the Elizabeth Thompson Science Fund is gratefully acknowledged. The erythroidine was generously supplied by Merck & Company, Rahway, New Jersey. Some of these results were reported at the meeting of the American Physiological Society in Boston, April 3, 1942.

[1] For a visual comparison of the behavior of these two preparations, decorticate and dissociative, *cf. Views of a decorticate dog* (C. H. Stoelting Company, Chicago), supplementary to a written report (**9**), and *The psychodynamics of dissociated behavior in the drug-state.*

Procedure

The experimental procedure has been described in its essentials elsewhere (**6**) and only slight modifications were introduced to include the recording of *BP*. Under ether anesthesia, the common carotid artery was located adjacent to the trachea before its division into the external and internal branches. The artery was separated from the adjoining vagus nerve by careful removal of the connecting fascia. A glass cannula was inserted into the artery just posterior to the larynx and then connected to a mercury recording manometer with high-pressure rubber tubing. The entire system, from tip of cannula to mercury column, contained 5 percent sodium citrate solution which prevented clotting of the blood. A nichrome-wire hot-point needle (heated by 1½ volt dry-cell) was attached to the float resting in the free arm of the mercury manometer and the fluctuations in *BP* were recorded directly on wax paper fed by a Parker continuous polygraph. The time was recorded in seconds on the polygraph by means of a Harvard clock and a third marker recorded successive presentations of the stimuli.

When these preparations were completed, dihydro-*B*-erythroidine hydrobromide (1 percent distilled water solution) was injected intravenously, artificial respiration was induced by tracheal insufflation (Harvard respirator) and functional tests were initiated. Conditioning was established by paired presentations of 250-watt spotlight (containing focusing lens) flashed into the animal's eye for four seconds (*L*) and terminated with the simultaneous presentation of a momentary shock (*US*) to paw of foreleg. On test-trials *L* would be given without *US*. Five dogs, ranging in age from 2 to 4 months, were used, and the following tests were made: (1) *R 1* and *R 2*—diffuse *CR* were developed under erythroidine and followed by *BP* tests some three weeks later; (2) *R 3* tested under erythroidine to determine the effect of light upon *BP* in the naïve animal; (3) *R 4* and *R 5*—first tested in the naïve state (like *R 3*) then conditioned to light, changes in *BP* being recorded throughout.

Results

R 1.—This dog was first conditioned in the erythroidine drug state by paired combinations of light (4 seconds) and shock applied momentarily to forepaw during the last half of the *L*-interval.[2] Three drug sessions were given on October 15, 1941, from 10:42 A.M. until 4:16 P.M. A total of 52 conditioning trials produced the following scores: 1st Session = 1/5 (5th trial), 5/5 and 9/9 (shock applied 13/19); 2nd Session = 10/13 (none with shock), 10/10 and 7/10 (shock given 17/20). The numerators of the above fractions indicate the number of *CR*'s recorded and the denominators, the total number of trials. In all, there were 42 *CR*'s out of a total of 52 trials, on only 30 of which shock-reinforcement was applied. The *CR* consisted of pupillary (*cf.* **7**) and striated

[2] Application of shock in this fashion was made in order to facilitate motion picture recording of the pupillary conditioned responses, which are to be reported in a separate paper (**7**).

muscle (*cf.* **6**) components. Prostigmin was used to hasten recovery and reduce the intervals between successive test-sessions.

On November 6, 22 days later, the left carotid artery was ligated, and the right carotid artery was cannulated and connected to *BP* recording apparatus. Erythroidine (20 mgm.) was then injected intravenously at 3:45 P.M. Upon the return of breathing reflexes at 4:34 P.M., a second dose of 40 mgm. was given at 4:36 P.M., followed by a final supplement of 40 mgm. at 5:10 P.M. As shown in Table I, the *CR* also contained a blood pressure (*BP*) component, since there was an increase in *BP* on each of the first five successive presentations of *CS* (*L*-1 to *L*-5). With suc-

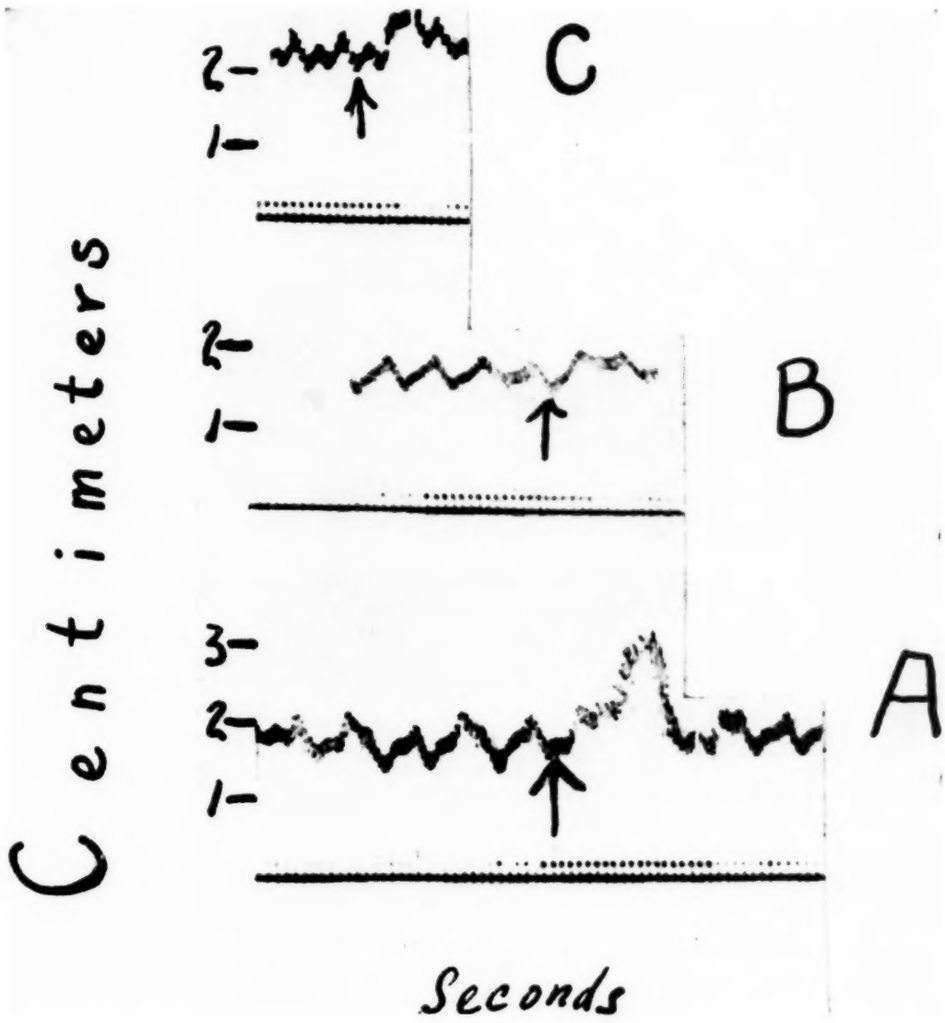


FIG. 1. Onset of Light is indicated by arrow (↑). *A* = *L*-1 is the initial test 22 days after original training. *B* = *L*-5 is the fifth successive presentation of Light, 5½ minutes later. Eleven subsequent light-presentations were completely negative. *C* = *BP–CR* after 6 reconditioning trials (also *cf.* Table I and text). The minute vertical striations are the pulse recordings and the more or less regular rise and fall is due to respiration.

cessive trials, however, the magnitude of the *BP–CR* decreased (40 mm. on the first to 14 mm. on the fifth). These figures, of course, are only the peak values and not truly representative of the total change in *BP*. For illustrative purposes, exact replicas of the polygraph recordings of the first and fifth trials are presented in Fig. 1, *A* and *B*. These tests, without shock reinforcement, resulted in complete extinction, and no change in *BP* was detected in the following 11 trials (*L*-6 to *L*-18). During these latter tests, from 4:34 P.M. until 5:01 P.M., shock-alone was applied on 12 different occasions. The application of the shock each time produced a considerable rise in *BP*, varying in magnitude from 20 mm. to 60 mm. (not included in Table I). These tests

TABLE I *
R 1

BLOOD PRESSURE CONCOMITANTS WITH DIFFUSE *CR* TO LIGHT IN DRUG-STATE

Time	*L*	*L* + Sh.	*CR* to L(ight)		*UR* to Sh(ock)	
			BP in mm.	Latency in secs.	*BP* in mm.	Latency in secs.
3:52	1		40	10		
	2		16	3		
	3		20	8		
	4		20	5		
3:56½	5		14	3		
3:57	6		—			
4:34	7		—			
5:01	18		—			
5:16		1			—	
½		2			14	6
5:17		3			20	4
¼		4	−10	3	10	1
½		5			4	3
¾		6			10	4
5:18¼	1		14	4		
½	2		10	2		
¾		7			14	1
5:19½		8	−20	4	16	1
5:20¼		9	−26	4	12	5
¾		10			14	3
5:21¾		11			10	2
5:22¼	3		14	5		
¾	4		12	4		

* Interpretation: *L* = Light-alone; *L* + Sh. (*CR* trial) = Light with terminal momentary shock applied to forepaw. *CR* to *L* = increase in Blood Pressure (*BP*) either with *L*—alone or during *CR* trial before shock was applied. Change in *BP* was recorded in millimeters (mm). Scores prefixed by minus (−) sign indicate drop in *BP*. Latency is in terms of the specific stimulus. For example, *L*-1 (3:52 P.M.) produced a peak rise of 40 mm. in *PB* ten seconds after *the start of L*. On Re*CR* no. 2 (5:16½ P.M.), no *CR* was evoked (during the presentation of *L*) but *UR* to shock was a rise of 14 mm, 6 seconds *after application of Sh*. Again, on Re*CR* no 8, there was a drop (−) of 20 mm in *BP* to *L* on fourth sec. just before onset of Sh., followed by a rise of 16 mm produced by Sh. (one sec. after shock was given). No *UR* scores are shown for *L*-trials since Sh. was not applied on these tests.

offered further proof that the change in *BP* initially evoked by Light (*L*-1 to *L*-5) truly represented a conditioned reaction, since the shock-tests after extinction failed to result in *pseudo* conditioning. At 5:16 P.M., reconditioning (re *CR*) was initiated in the more conventional manner: light (4 seconds) terminated with the simultaneous presentation of momentary shock. After 6 training trials, the *BP*–*CR* reappeared on *L*-1 as a sustained rise of about 14 mm., with a latency of 4 seconds (*cf.* Table I at 5:18¼ P.M., and Fig. 1-*C*).

It is significant that when the trials were spaced too closely together (15-second intervals) the magnitude of the *UR* to shock is smaller (4, 5, and 6). Since a maximum *BP*-increase produced by *US* cannot be further increased by a second closely following shock-stimulus, sufficient time must elapse between trials to permit the *BP* to return to pre-shock level. When this precaution is taken, by increasing the time interval between successive trials (7 to 10), the *UR* is stronger. The significance of the momentary fall in *BP* on some (4, 8, and 9) is not entirely clear. It is undoubtedly part of the conditioning process since it has never been found in the untrained animal, but its exact role is yet to be determined. Prostigmin was given at 5:24 P.M. to counteract

the effect of the erythroidine. After a suitable interval, light was presented, but the *CR* was completely repressed.

R 2.—This animal had been used in several experiments before the tests were made. These earlier procedures and results are largely irrelevant here and will be referred to only when necessary. On October 6 and 7 the animal was subjected to 6 sessions with erythroidine during which generalized (striated muscle) *CR* to tone was well established in the usual way (*cf.* **6**) and recorded on motion picture film. On October 23 the right vagus nerve was completely cut at the level of the larynx, under ether anesthesia. From October 24 through November 1 the animal was

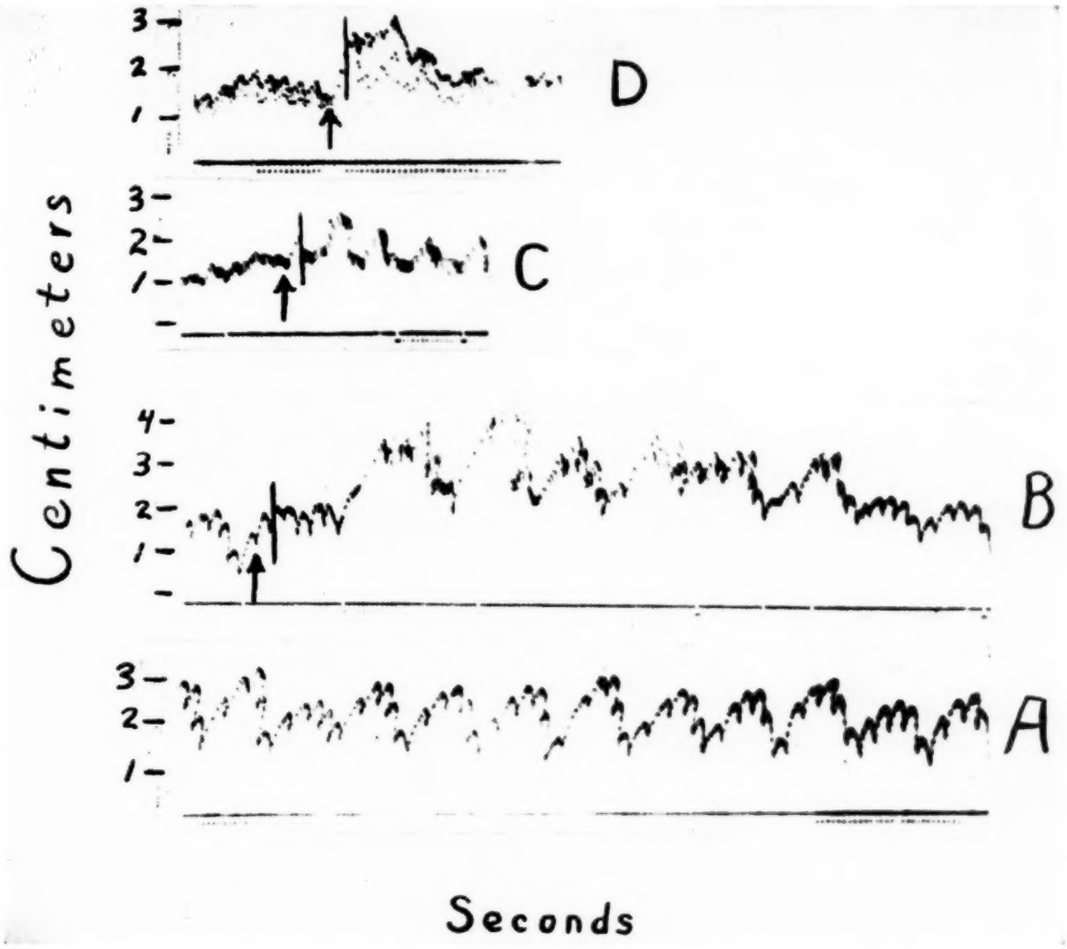


FIG. 2. *A* = cyclic *BP* time—*CR*. *B* = 1st (re)conditioning trial at 2:16 P.M. *C* = *L*-1, three minutes later (no shock). *D* = *L*-14 (no shock). Arrow indicates onset and vertical line termination of Light. On training (*CR*) trials, momentary shock was applied with termination of *L*. On test-trials, Light-alone (*L*), shock was omitted. *Cf.* text for interpretation.

subjected to 6 more drug-sessions during which the diffuse-generalized-*CR* was now developed to a light-stimulus and again recorded in motion pictures. In all, at least 200 shock-reinforced trials were given at 15-second intervals. During intervening normal periods, the *CR* was completely suppressed; the animal acted like a naïve subject, remaining completely indifferent to *CS*.

On November 13, under nembutal anesthesia, the visual cortex was completely removed from both hemispheres. On November 27, the animal was tested in the *normal state*. Upon application of stimulus at 1:10 P.M., "the animal began to tremble and withdraw from the light" (from the protocols). At 1:25 P.M., a dose of 20 mgm. of erythroidine was injected (13th session) and, under the drug, at 1:30 P.M., Light produced a diffuse *CR*, consisting of attenuated twitches of the striated muscles of the entire body. Due to the myoneural block, the quality and magni-

tude of the striated muscle response in the drug-state was different from the *CR* evinced in the previous normal period.[3]

The animal was then prepared for *BP* tests in the usual way. The right carotid artery, being quite flabby, was ligated and *BP* was recorded from the left artery. As shown in Fig. 2 *A*, the form of the *BP* recordings was unique. The usual pulse and respiratory variations were accompanied by a third wave of largest amplitude and time sequence, approximating a 15-second time interval. Although a slow *BP* wave has been noted (*cf. No. R 3* below), this type of cyclic rise and fall approximating a 15-second interval has never been observed in the untrained animal.

Eight trials with Light-alone (2:07 P.M. to 2:15 P.M.) had no effect on this phenomenon. At 2:16 P.M., one *CR* trial was given in the usual way. Again, *L* produced no effect but shock evoked a marked rise in *BP* (Fig. 2 *B*). Three minutes later, a clearcut *CR* was evoked by Light-alone: a peak rise of 10 mm., with a latency of 4 seconds (Fig. 3 *C*). The *BP–CR* to *L* was thus reinstated. The clear-cut periodic time-cycle was disrupted, although, with continued training, there was some tendency for it to reappear. The results are summarized in Table II.

TABLE II *
R 2

BLOOD PRESSURE CONCOMITANTS WITH DIFFUSE *CR* TO LIGHT IN DRUG-STATE

Time	*L*	*L* + Sh.	*CR* to L(ight)		*UR* to Sh(ock)	
			BP in mm.	Latency in secs.	*BP* in mm.	Latency in secs.
2:16		1			30	22
2:19	1		10	4		
2:20	2		16	6		
2:32		12			12	5
2:33		13	10	4	32	5
2:34	6		—			
½	7		8	6		
2:35	8		8	4		
2:36	9		6	4		
2:37	10		—			
½		14	6	4	10	1
2:38		15			—	
2:39		16	−8	2	30	8
2:41	11		—			
½	12		—			
2:42		17	10	4	46	5
2:43		18	20	4	24	3
2:44	13		20	5		
2:46	14		26	4		

* For explanation *cf.* Table I and text.

R 3.—On November 7, the animal was prepared for *BP* recording in the usual way and then injected with erythroidine. This untrained animal was tested from 4:25 P.M. to 5:00 P.M. under the drug with 14 spaced trials of Light-alone. All trials were completely negative, producing no detectable changes in *BP*. Clearly, the *BP* noted in *R 1* and *R 2* was the result of the conditioning

[3] These observations confirmed the data previously reported concerning the effect of the removal of the temporal lobes upon auditory *CR* developed in the drug-state (**5**). As a result of extirpation of the cortex, conditioning in both normal and drug-state was established through subcortical pathways. This prevented or disrupted the drug-dissociation (depending upon the time of surgical intervention) so that the *CR* developed in one state now appeared spontaneously in the other. In the present case, the drug-state *CR* was maintained for a period of 26 days without interpolated training. During this period, the dissociation was disrupted by the removal of the visual areas of the brain. The drug-state *CR* thereafter appeared in the normal state as well, since cortical inhibition had been removed.

process, since in the present untrained animal, *L* was completely ineffectual in producing such changes in *BP*. The first two and the last three trials of this series in *R 3* are reproduced in Fig. 3. The minute vertical striations (clearer in *L*-1 and *L*-2) were the pulse recordings. The second rise and fall, averaging about 3 seconds per complete cycle (clearer in *L*-12 to *L*-14) was due to the respiration. The significance of the third slow wave occurring at about 25-second intervals in *L*-12 to *L*-14 is not at all clear. Shock was never applied to this animal, and therefore this slow curve was *not* a time-*CR* resulting from the experimental procedure.[4]

R 4.—This naïve animal was tested on November 10 and the results are summarized in Table III. Erythroidine (20 mgm.) was injected at 12:14 P.M. While slight breathing reflexes were noted at 12:40 P.M., a supplementary dose was not required until 1:00 P.M., at which time 20 mgm. were injected.

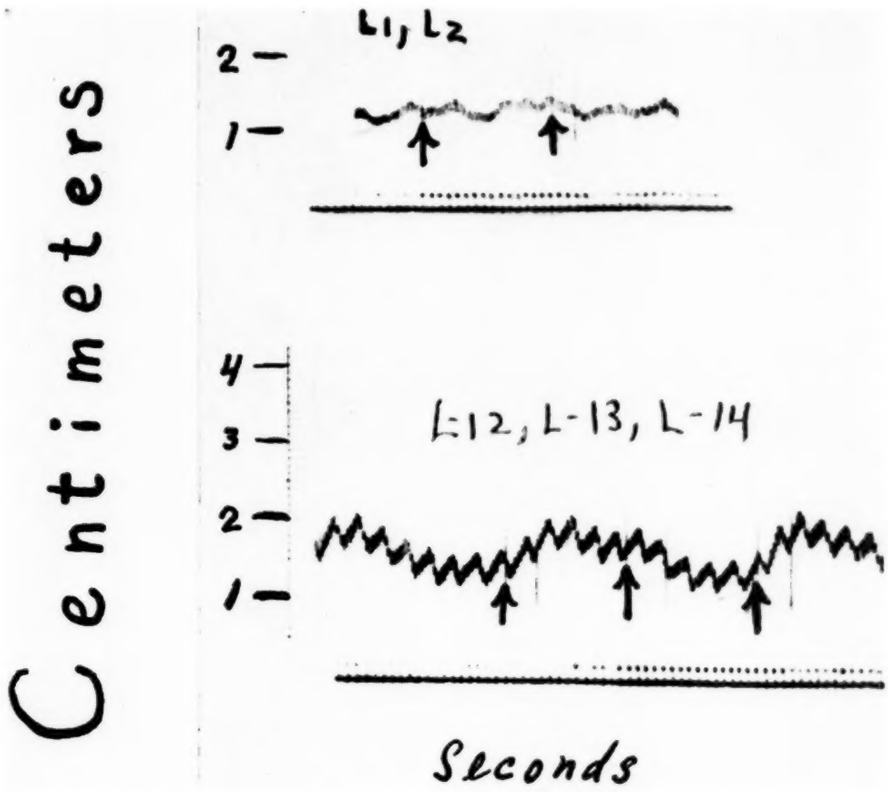


FIG. 3. Arrow indicates onset and vertical line termination of Light-stimulus. For interpretation, *cf.* text.

CR training was initiated after preliminary tests indicated that Light was negative. Since the first five training trials were spaced at only 15-second intervals (*cf.* Table III and Fig. 4), there was some tendency for the *UR* to be suppressed (also *cf. R 1*). With better spacing, the *UR* appeared every time (trials 6 to 8). The first *CR* to *L*-alone occurred after 8 training-trials. A maximum *BP–CR* of 22 mm. occurred with *L*-4, after 2 more training-trials. Replicas of the polygraph recordings of these two tests and the first six training-trials are shown in Fig. 4. The absence of *CR* on *L*-7 and *L*-8 was due to a temporary depression which followed the supplementary drug-injection at 1:00 P.M.[5] After a short pause, the *CR* reappeared on *CR*-13 before shock was applied.

[4] Such waves apparently result from a number of internal conditions affecting the sympathetic nervous system. We have observed them in several animals but they have never coincided with the time interval of the stimuli presentations *prior* to conditioning. When light and shock combinations have been presented successively at 15-second intervals for a number of trials, the blood pressure has been noted to rise and fall at approximately this time interval (*cf. R 2*).

[5] Following injections of the drug, there was always a temporary drop in *BP*, varying from 20 mm. to 100 mm., which persisted for a couple of minutes.

TABLE III *
R 4

DEVELOPMENT OF BLOOD PRESSURE *CR* IN NAÏVE ANIMAL IN DRUG-STATE

Time	*L*	*L* + Sh.	*CR* to L(ight)		*UR* to Sh(ock)	
			BP in mm.	Latency in secs.	*BP* in mm.	Latency in secs.
12:38	1–4		—			
41			Shock alone		12	7
41¾			" "		24	10
42		1			—	
42¼		2			26	5
42½		3			—	
42¾		4			12	2
43		5			—	
43½		6			16	6
44		7			14	4
44½		8			10	8
45	1		16	11		
	2		—			
	3		—			
46		9	6	2	30	5
		10	8	4	20	8
48	4		22	8		
51	5		10	6		
52		11			35	12
		12	4	3	40	16
53	6		16	5		
1:01	7		—			
02	8		—			
04		13	2	4	14	5
05		14	9	2	32	10

* For explanation, *cf.* Table I and text.

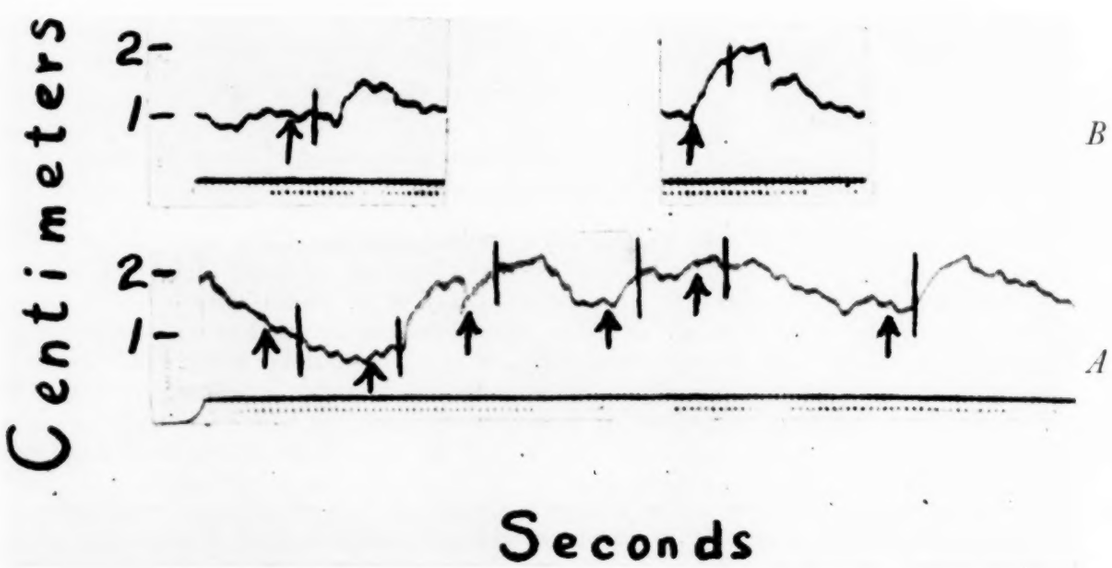


FIG. 4. *A* = *CR* training trials, no. 1 to no. 6 inclusive. *B:* Left to right, two test trials without shock: *L*-1, after eighth *CR* trial and *L*-4 after tenth *CR* trial. In latter case, note anticipatory rise in *BP* to *L* before momentary shock would have been applied. Also *cf.* Figs. 1–2 and text.

R 5.—This naïve animal was tested on November 11. Some difficulty was encountered in recording from the right carotid artery which was therefore ligated. The left carotid artery was then cannulated and connected to the recording system in the usual way. A dose of 20 mgm. of erythroidine was given at 4:35 P.M. and a supplementary dose of the same size was injected at 5:56 P.M. As indicated in Table IV, the results confirmed the data of *R 4*. *BP* failed to change

TABLE IV *
R 5

DEVELOPMENT OF BLOOD PRESSURE *CR* IN NAÏVE ANIMAL IN DRUG-STATE

Time	*L*	*L* + Sh.	*CR* to L(ight)		*UR* to Sh(ock)	
			BP in mm.	Latency in secs.	*BP* in mm.	Latency in secs.
4:41	1–4		—			
49		14			24	13
5:06	8		—			
10		21	4	3	22	8
½	9		4	3		
¾		22			26	8
11½		23	6	4	16	6
14½	10		10	6		
15		24	6	5	22	6
21		28	4	4	18	9
22	14		12	7		
26	16		6	3		
28		29	2	5	20	10
30	18		6	4		
31		31	2	2	20	8
32	19		10	8		
35		32	4	4	24	9
36½	22		18	5		
42	32		6	7		
43	33		4	6		

* For explanation *cf.* Table I and text.

to *L* before training, whereas clear-cut increases from 12 mm. to 18 mm. were recorded after conditioning was established.

Definite breathing reflexes were not noted until 5:21 P.M., which were the only overt manifestations until striated muscle reactions were evoked by shock seven minutes later.[6] The blood pressure *CR* was therefore developed before overt reactions were possible. Polygraph recordings of representative trials are shown in Fig. 5.

It is to be noted that the *CR* developed more slowly in this animal. The first *CR* did not occur until the 21st training trial, whereas it first appeared in *R 4* after 8 trials. This may be due

[6] Voluntary breathing of course is to be distinguished from breathing reflexes. If artificial respiration were discontinued upon the appearance of breathing reflexes, death would shortly ensue. Breathing reflexes normally appear some time before striated muscle reactions are evoked as *UR* to shock.

[7] It is almost impossible to keep these factors entirely constant, since depth of dose is largely empirical and supplementary injections are required when test-time is increased by delays produced by clogging of cannulas, etc. Furthermore, it is extremely difficult to ascertain the proper dose in the first sessions.

Certain other facts are relevant. The conditioning process may have been impeded in *R 5* due to the fact that the carotid supply to the brain was entirely cut off. In *R 4* one carotid pathway was still intact. It is true that the entire carotid supply to the brain was also eliminated in *R 1* and *R 2*. However this was done only *after* the initial conditioning had already been well established in both of these animals.

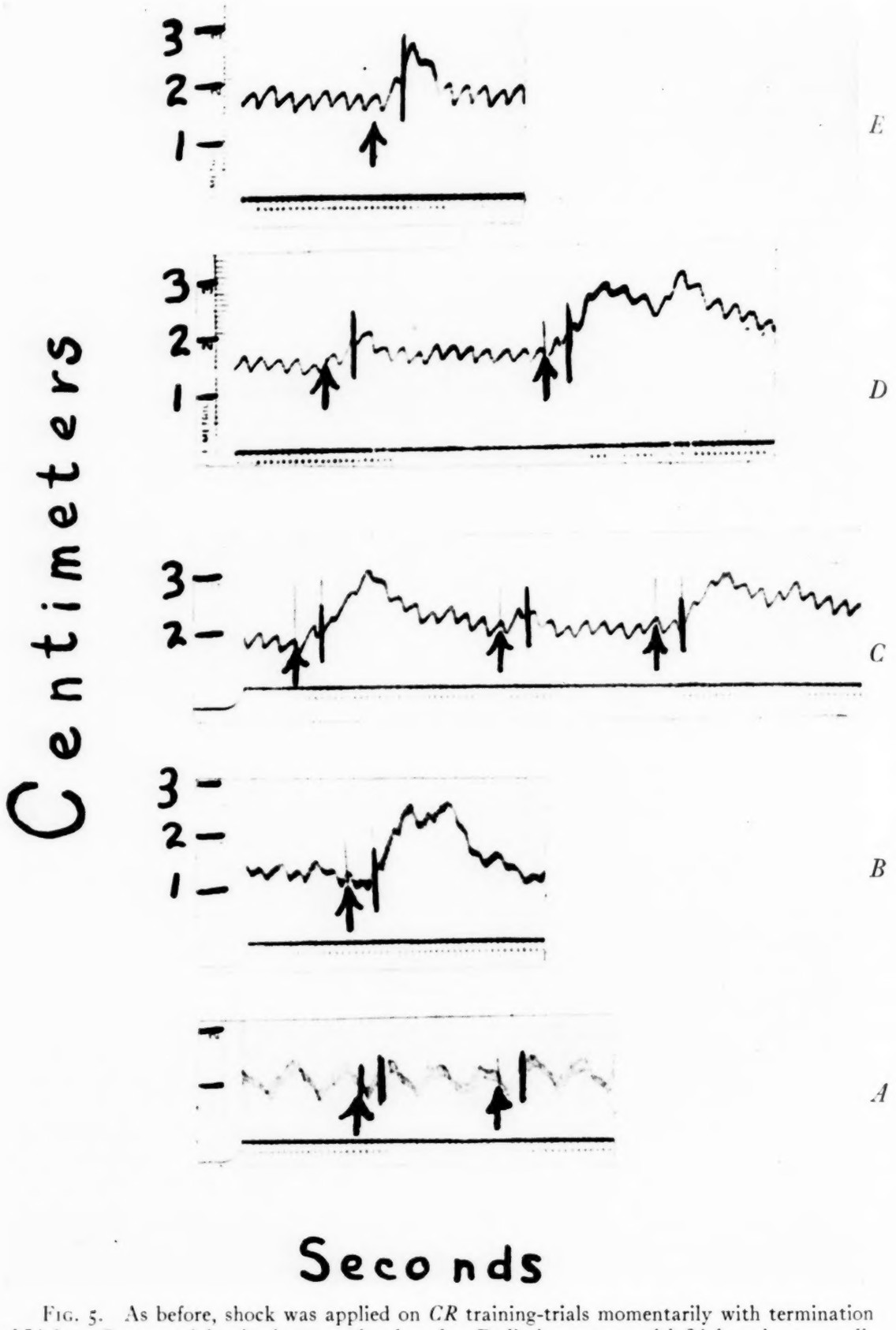


FIG. 5. As before, shock was applied on *CR* training-trials momentarily with termination of Light. On test trials, shock was omitted. *A* = Preliminary tests with Light prior to conditioning (*L*-1 and *L*-2). *B* = 14th *CR* training-trial. *C:* From left to right, *CR* no. 21, *L*-9 and *CR* no. 22. *D:* In same order, *L*-10 followed by *CR* no. 24. In latter, note anticipatory rise in *BP* prior to application of shock. *E* = *L*-22. Again there is an anticipatory rise in *BP* before momentary shock would have been applied.

to individual differences or a greater depth of the drug-state in the present case.[7] Support for the latter interpretation is derived from the protocols in which it was recorded that during the earliest trials, starting at 4:43 P.M., "Pupil fails to constrict to light and only direct response of muscle was evoked by shock."

A check on the validity of the *CR* was made by attempting to extinguish it by successive presentations of light without shock (*L*-22 to *L*-33). Although this was not completely successful, the *BP–CR* was reduced to a barely detectable minimum, *L*-33 evoking a rise of 4 mm. Between extinction trials *L*-32 and *L*-33, the apparatus was activated in the usual way but the light was omitted. No change whatsoever in *BP* occurred. On the next trial *L*-33, the minimum *CR* noted above was detected.

Discussion

In the drug-state induced with erythroidine, the conditioning of blood pressure (*BP*) proceeds uneventfully, is susceptible to extinction and can be reconditioned. One cannot fail to be impressed by the relative ease with which the *BP–CR* is established. In the poorest subject, *R* 5, the *CR* appeared on the 21st trial and was present thereafter. In *R* 4 the *CR* appeared after 8 training trials. Under ideal conditions of dosage and *BP* recording, it should be possible to establish this learning within a trial or two. In data published on diffuse *CR* in the drug-state, it has been shown that retention has been maintained from 5 to 6 days, without intervening training (**6**). The retention of *BP–CR* is even more impressive. With but a total of 30 original (shock) training trials in three successive sessions on a single day, the *BP–CR* was still present in *R* 1 after a lapse of more than three weeks.

The rapid development and prolonged retention of the drug-state *CR* is due to several factors, one of which seems to be the lack of other, normal, activities which would impede its development. Contrast these findings with the results of conditioning localized motor responses in normal dogs. In 17 normal animals, the number of training trials required to procure localized paw-flexion (86 percent average efficiency) ranged from 100 to 250 (**3**, p. 300). On the other hand, isolated semitendinosus muscle conditioning developed under curare after 23 to 125 trials (100 percent efficiency) (**8**, p. 266).

During the training in the normal situation, learning is handicapped by interference from other psychological activities. In the drug-state, however, besides the basic physiological processes, there is nothing else occurring internally other than the psychosomatic activity produced by the conditioning stimuli. These results agree with the well known investigations which have shown that there is no loss in the retention of nonsense syllables during sleep. "Forgetting is not so much a matter of decay of old impressions and associations as it is a matter of interference, inhibition, or obliteration of the old by the new" (**10**, p. 612).

In the present study *BP* conditioning was developed in the deep

drug-state before breathing reflexes or striated muscle reactions appeared. Similar results were procured when the animals were capable of these overt responses. From these data it seems clear, therefore, that conditioning in the drug-state, whether of the isolated muscle (**8**) or of the diffuse struggle pattern (**6**), involved a blood pressure component.

As shown by the results of the previous studies with erythroidine and curare, all elements of the drug-state *CR* are completely repressed upon recovery from the drug (**2, 6, 8**). Only occasional tests were made specifically with *BP* after the injection of prostigmin (*cf. R 1*) and systematic recording of *BP* in the normal state by the present technic has not been made for obvious reasons. It is safe to infer, however, that the *BP–CR*, developed in the drug-state, was suppressed after recovery as were the other components of the drug-state *CR* which have been studied. The behavioral picture confirmed this conclusion. The normal animal remained unaffected by the *CS* which evoked no conditioned startle or respiratory reaction nor any change in pulse rate or body stance. The animal behaved, to all intents and purposes, like a naïve subject. The dissociation was so complete that no 'defense' *CR* developed to the experimental situation. Truly the animals were *amnesic* for the drug-state experiences [*cf.* (**6**) for elaboration of this point].

The present data have implications for the more conventional conditioning and learning experiments. It has been suggested elsewhere (**4**) that it was a mistake to extend the Pavlovian doctrine of constant reinforcement from the salivary to motor conditioning technic. Whereas this procedure will make for a stronger *CR* in the former, it will in the latter case tend to prevent the proper development of conditioning. [Applying the shock-stimulus on every trial with the *CS*, whether the animal flexes the paw or not, will not tend to strengthen paw-flexion. To the contrary, this procedure simultaneously tends to develop paw-*extension*. The reason becomes clear in the following analysis. Suppose *CS*, light or bell, is presented and the animal lifts the paw. Shock is now applied when the paw is in flexed position. The reciprocial response is now the only possible reaction: namely, extension. In conditioning terminology, application of shock to flexed paw really constitutes an unconditioned stimulus for paw extension.] From this reinforcement of opposed responses, conflict and extinction would naturally result. In other words, *rewarding with punishment* will tend to 'stamp out' the desired response (paw-flexion). Whatever terminology one prefers, psychologically one must expect unstable conditioning under these circumstances, which is what is reported by those attempting to apply the Pavlovian dictum to motor conditioning (**11**).

This analysis holds not only for the total organismic situation, but the single muscle as well. If shock is applied to the isolated semitendinosus muscle when it is contracted, if anything happens at all, the muscle *relaxes* (*cf.* **1**, p. 140). The present data offer further confirmation for this argument. If the *CR* trials are too closely spaced together, the first shock produces a *BP* rise as is expected. But there is no further rise in *BP* upon application of second shock following too closely after the first one. This is analogous to applying reinforcement when the desired response is already evoked. This error was made in the early stages of the present study because of the practice of applying stimuli at intervals of about 15 seconds in motor-conditioning training in the normal situation. Thus when the trials for *R 4* were spaced but 15 seconds apart, the *BP* failed to rise to shock on the alternate trials. When the interval was lengthened to 30 seconds, the *BP* response was procured on every trial (*cf.* Table III). The continued use of a short interval between successive presentation of the stimuli will, of course, result in a time-*CR* (*cf. R 2*) such that the response will now be made at regular intervals and the conditional stimuli used by the experimenter may be ineffective. This is to be expected since the *real CS* is now the time interval between the successive presentations. Apparently the internal time-mechanism may be more prepotent than external stimuli.

(Manuscript received January 19, 1942)

References

1. Culler, E. A. Recent advances in some concepts of conditioning. *Psychol. Rev.*, 1938, **45**, 134–153.
2. Culler, E. A., Coakely, J. D., Shurrager, P. S., & Ades, H. W. Differential effects of curare upon higher and lower levels of the central nervous system. *Amer. J. Psychol.*, 1939, **52**, 266–273.
3. Culler, E. A., & Mettler, F. A. Conditioned behavior in a decorticate dog. *J. comp. Psychol.*, 1934, **18**, 291–303.
4. Girden, E. Conditioning and problem solving behavior. *Amer. J. Psychol.*, 1938, **51**, 677–686.
5. Girden, E. Cerebral mechanisms in conditioning under curare. *Amer. J. Psychol.*, 1940, **53**, 397–406.
6. Girden, E. Generalized conditioned responses under curare and erythroidine. *J. exp. Psychol.*, 1942, **31**, 105–119.
7. Girden, E. The dissociation of pupillary conditioned reflexes under erythroidine and curare, *J. exp. Psychol.* (In press.)
8. Girden, E., & Culler, E. A. Conditioned responses in curarized muscle in dogs. *J. comp. Psychol.*, 1937, **23**, 261–274.
9. Girden, E., Mettler, F. A., Finch, G., & Culler, E. A. Conditioned responses in a decorticate dog to acoustic, thermal, and tactile stimulation. *J. comp. Psychol.*, 1936, **21**, 367–385.
10. Jenkins, J. G., & Dallenbach, K. M. Obliviscence during sleeping and waking. *Amer. J. Psychol.*, 1924, **35**, 605–612.
11. Liddell, H. S. The conditioned reflex. *In* Moss, F. A. (editor), *Comparative psychology*, New York: Prentice-Hall, 1934.

THE PHENOMENON OF THE MUSCLE-TWITCH IN FLEXION CONDITIONING[1]

BY N. H. PRONKO AND W. N. KELLOGG

Indiana University

In the conditioning of the knee-jerk with human subjects, Wendt (4) has reported the appearance of two distinct kinds of conditioned responses. His method was to condition the extensor muscles of one leg to a tap on the patellar tendon of the leg on the opposite side of the body. Both conditioned and unconditioned stimuli were therefore in the same sense modality. Concerning the two sorts of *CR*s which were noted, Wendt has written (4, p. 92): "Repeated successive stimulation by two blows is found to result in two kinds of conditioned responses. . . . One kind of response is a conditioned kick of the second responding leg on the omission of the second blow. It develops only after prolonged double stimulation and is easily extinguished. Its latency is typically 200 to 300 sigma." The latency of the other response was 120 to 180 sigma.

The response of shorter latency, Wendt thought, should be classed as 'an experimentally produced crossed reflex.' It appeared much earlier in the training and depended upon 'an original integration of stimulus and response' (p. 83). The response of long latency, on the other hand, he considered as a 'true' conditioned response.

Phenomena similar to these two sorts of reactions have recently been observed in the Indiana laboratory during the flexion conditioning of dogs. The two types of *CR*s have been found to occur, moreover, almost exclusively in cases where both conditioned and unconditioned stimuli are in the same sense modality, and but very rarely with the traditional buzz-shock conditioning technique.

METHOD

The method by which the two *CR*s have been most readily observed may be called the shock-shock conditioning procedure. In our laboratory this has consisted simply in substituting an additional electric shock for the customary buzz of the buzz-shock method—without changing the unconditioned stimulus or the duration or spacing of either of the stimuli. The conditioned stimulus has consisted of a 1000 ~ *alternating current* shock of 2 sec. duration delivered preferably to one of the forepaws. The unconditioned stimulus has consisted of the usual make-break *direct current* shock, .2 sec. in length, which occurs simultaneously with the last .2 sec. of the conditioned

[1] The observations reported in this communication from the Indiana Conditioning Laboratory were made possible by a series of grants-in-aid of research from the University and by continual assistance from the National Youth Administration of the United States Government.

alternating-current shock. If the conditioned shock is sent to one of the front feet, then the unconditioned shock is sent to one of the rear feet. Both stimuli are delivered through stainless steel discs 2 cm. in diameter which are securely taped to the skin at the points desired. The hair of the animal is removed by clipping in these areas.

We have found it best to control the intensity of the two stimuli according to the behavior of the subject. The unconditioned stimulus has customarily been adjusted so as to produce a flexion reflex, in the foot to be conditioned, of 4 in. in extent. In order to insure the reception of the conditioned stimulus, it too has been adjusted to produce a sustained lift of the receptive foot for the entire 2-sec. period of its duration. The animal is supported by slings under each leg and by a hammock under the chest and abdomen so that he can lift any one or all of his feet at any time without falling.

Other details of the procedure are identical with those of the buzz-shock method, a full description of which has been published elsewhere (2).

Results

In this experimental situation most dogs develop a flexing twitch or slight jerk of the rear foot, which occurs almost simultaneously with the response of the front foot to the conditioned alternating-current shock. The first appearance of this muscle-twitch is early in the training although it does not usually become regular or consistent till later. It is possible that it starts as a kind of spread or generalization of the front-foot response. In some instances the muscle-twitch is observable not only in the member to be conditioned, but in the other legs as well. However, such a generalization is not likely to persist for a long period. After a few pairings of conditioned and unconditioned stimuli, the twitch response is usually confined to the limb to which the unconditioned stimulus is sent. For convenience we have called the muscle-twitch reaction a type-T conditioned response, and the *bona fide* or ordinary conditioned flexion reaction a type-L conditioned response. There are some grounds for thinking of the type-T reaction as a conditioned *reflex* and the type-L reaction as a conditioned *response.*

In confirmation of the observations of Wendt (4), we find the latency of the muscle-twitch (type-T) *CR* to be short—less than 200 sigma. With the present timing of conditioned and unconditioned stimuli, the latency of the conventional flexion *CR* (type-L) is usually between 400 to 1200 sigma. The amplitude of the type-T *CR* is also small—.25 in. or less—and it remains practically constant throughout a long period of training. The amplitude of the type-L *CR*, on the other hand, starts at a low value and gradually builds up to a maximum of 2 in. or more in our experimental situation. The course of its development follows the typical sigmoid conditioning curve. The frequency of the type-T *CR* does, however, increase with training, but not exactly in the same manner as the frequency of the type-L *CR* (*vide infra*).

The patterns of the two responses are indicated in Figs. 1, 2, 3, and 4 which are reproductions of kymographic records of flexion

conditioning by the shock-shock method. The top line in all four figures is the record of movement of the left-rear foot. The line second from the top is the tracing of the right-rear foot which in the cases shown was the foot to be conditioned, *i.e.*, the foot to which

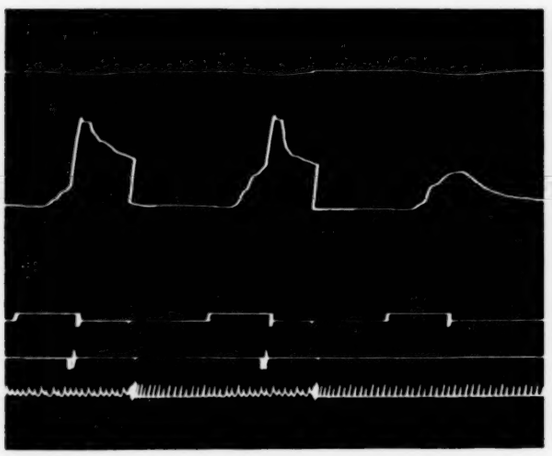


FIG. 1.

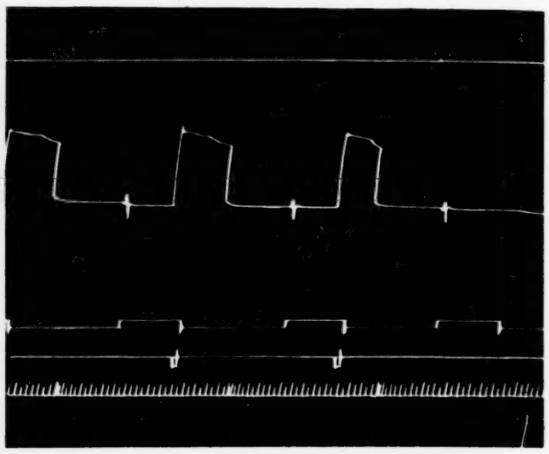


FIG. 2

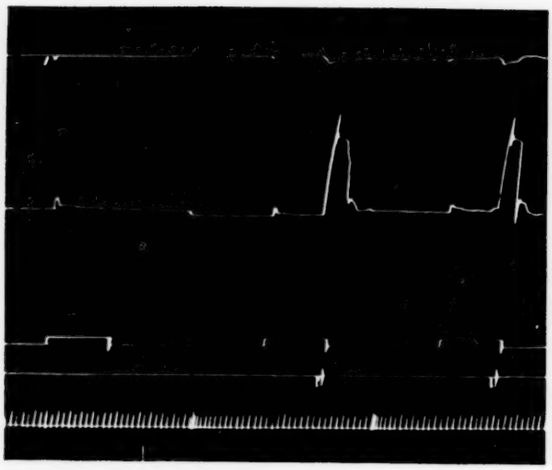


FIG. 3

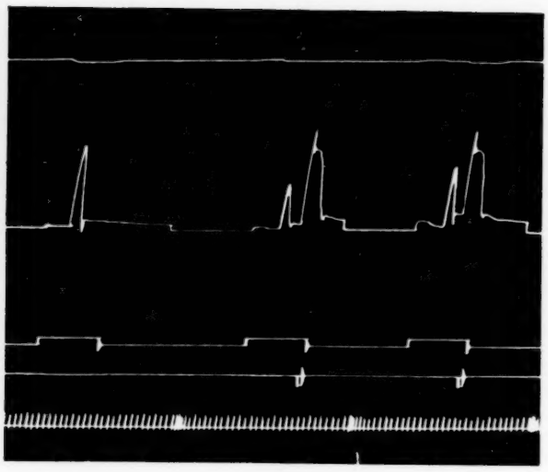


FIG. 4

Figures 1, 2, 3, and 4 are reproductions from kymographic records of type-T and type-L conditioned reactions obtained during the flexion conditioning of dogs. In all four figures, the top line traces the movements of the left-rear foot and the line second from the top is the record of the right-rear (conditioned) member. The last three lines are in order from top to bottom (1) the alternating-current conditioned stimulus (2) the d.c. unconditioned stimulus, and (3) time in units of .2 sec.

Figure 1 shows regular type-L *CR*s in the right-rear foot, and Fig. 2 shows contrasting instances of 'pure' type-T *CR*s. Figures 3 and 4 show combinations of type-T and type-L conditioned responses. (See text for further description.)

the unconditioned shock was sent. The line third from the top is the alternating-current, conditioned-stimulus line, and the last two lines show respectively the unconditioned shock, and time in units of .2 sec.

Figure 1 shows regular type-L *CR*s obtained by the shock-shock method. The two reactions to the left in Fig. 1 are combinations of conditioned and unconditioned responses. The irregular rise of the type-L *CR* is interrupted when the d.c. shock occurs, by the smooth arc of the unconditioned reflex. In the right-hand trial, the unconditioned shock has been omitted by *E* and the full contour of the type-L *CR*, uninterrupted by the unconditioned reflex, is pictured.

Figure 2 shows a contrasting instance of a 'pure' type-T *CR*. The latency is short, the amplitude is low, and the response ends almost immediately. That the type-T *CR* also occurs when the unconditioned stimulus is omitted, is shown in the trial to the right.

Figure 3 depicts a sort of combination of type-T and type-L responses. It will be noted in these instances that after the initial type-T lift, the leg is not immediately returned to its original position but remains slightly flexed or tensed. Yet even type-T movements such as these were seldom great enough to remove the foot completely from the floor. The change in the level of the response line, after the beginning of the type-L *CR*, indicates instead a postural shift which transfers the weight of the animal from one of the hind feet to the other. In the right-hand trial of Fig. 3, a further type-L lift occurs just before the beginning of the unconditioned reflex.

Figure 4 shows a more pronounced and obvious combination of the two kinds of conditioned reactions. The type-T *CR* of short latency, is followed about a second later by a well-defined type-L lift which, in this instance, reaches a height of from 2 to 4 inches. The completely conditioned nature of both type-T and type-L *CR*s

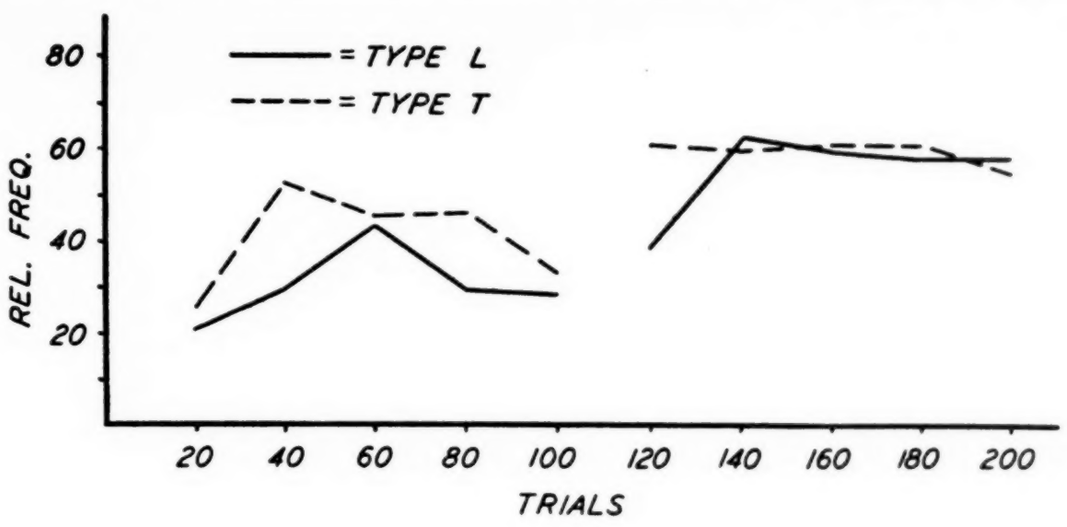


Figure 5 gives combined frequency curves (for 8 dogs) for type-T and type-L conditioned responses. The type-T curve is less like the usual sigmoid conditioning curve than the type-L curve.

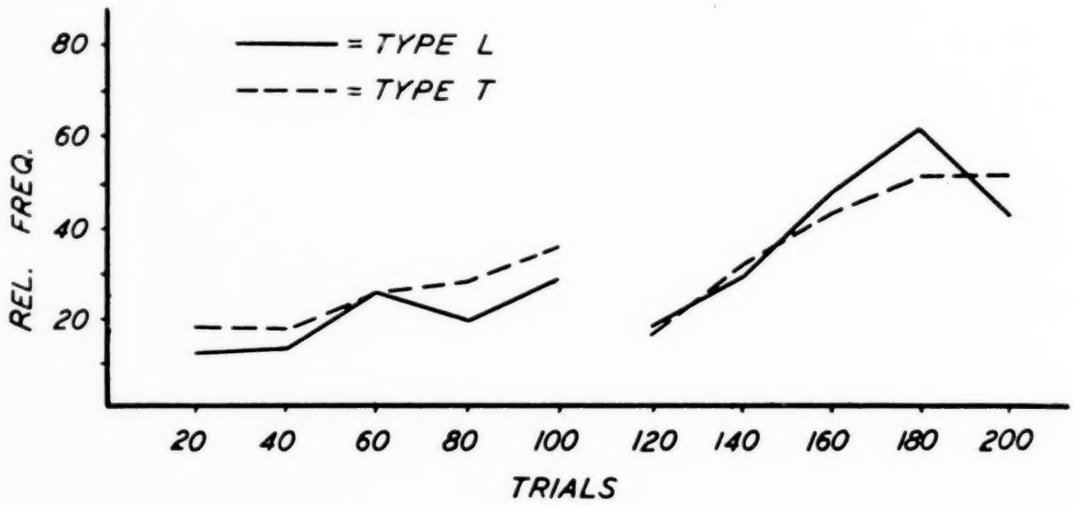


Figure 6 gives combined frequency curves (for 8 dogs) for type-T and type-L *CR*s obtained under the influence of nembutal.

is demonstrated in the trial to the left in which the unconditioned stimulus was omitted by *E*.

Additional differences between the two sorts of conditioned reactions are to be found in their different rates of acquisition—as shown in Figs. 5 and 6. The curves in these figures are composite relative frequency graphs for the conditioning of the right-rear foot in 8 animals. Figure 5 gives the frequency curves for 200 trials by the shock-shock method. Figure 6 shows comparable data for the same 8 animals when shock-shock conditioning was conducted while the subjects were under the influence of hypnotic doses of nembutal (pentobarbital sodium, Abbott). Each point on the curves is a frequency value obtained from averaging the results of 160 trials (20 trials for each of the 8 subjects). A suitable scheme of rotation, identical to that employed by Headlee and Kellogg (1), reduced to a minimum the effects of one type of training upon the other. The breaks in the graphs between points 100 and 120 represent the interpolation of a group of trials which are irrelevant to the curves as they are plotted.[2]

It will be seen from Fig. 5 that the curve for the type-L *CR*s starts at a lower level than the type-T curve, but reaches about the same height in the last group of 20 stimulations. It more nearly resembles the sigmoid conditioning curve than does the type-T graph, which begins at a higher level and which changes less throughout its length.

From Fig. 6 it appears that there is also a greater increase in the type-L *CR*s when training is conducted under the influence of

[2] For a full description of the scheme of rotation the reader is referred to Headlee and Kellogg (1).

nembutal. The depressing effect of the drug tends to lower the initial frequency of both curves so that the type-T frequency graph under the drug bears a closer resemblance to the usual conditioning curve than does the type-T graph of training in the normal state. It is evident from the results shown in Figs. 5 and 6, that the frequency of type-T *CR*s does increase with training. However, because of higher initial frequencies, the increase for type-T *CR*s is not as great as that for type-L *CR*s.[3]

Summary

When the flexion conditioning of dogs is elicited by the pairing of two electric shock stimuli, a kind of flexing jerk or muscle-twitch appears in the member to be conditioned shortly after the incidence of the conditioned stimulus. This twitch is distinct both from the regular flexion *CR* and from the unconditioned flexion reflex. The muscle-twitch response has here been designated as a type-T conditioned reaction, in contrast to the customary conditioned lift, which has been called a type-L conditioned reaction. There are some grounds for using the term 'conditioned reflex' to designate the type-T *CR* and 'conditioned response' for the type-L *CR*.

1. Both types of *CR*s are common during shock-shock conditioning. But the type-T *CR* has been observed only very rarely during buzz-shock conditioning. This may indicate that the continued elicitation of the type-T response (or conditioned 'reflex') requires the presentation of both conditioned and unconditioned stimuli in the same sense modality.

2. The type-T *CR* has a latency of less than 200 sigma. The type-L *CR* has a latency of from 400 to 1200 sigma.

3. The amplitude of the type-T *CR* is small and does not change much with prolonged training. The amplitude of the type-L response increases with training to a maximum of 2 in. or more.

4. The frequency of each kind of conditioned response increases with practice, but the frequency of the type-L *CR* increases more than that of the type-T *CR*. This is in part because the type-T *CR* begins with a higher frequency in the first place.

5. The two types of *CR*s seem to be somewhat akin to the two sorts of conditioned responses observed by Wendt during the conditioning of the knee-jerk. In Wendt's experiment both conditioned and unconditioned stimuli were also presented in the same sense modality.

[3] It has occurred to the present writers that the recently reported spinal conditioning of Shurrager and Culler (3) may in fact be type-T rather than type-L conditioning. The fact that these investigators conditioned a twitch of the semitendinosus muscle rather than a complete flexing movement of the limb suggests such a possibility, as does their method of producing the phenomenon, which was fundamentally that of shock-shock conditioning.

6. The suggestion is offered that the recently reported spinal conditioning of Shurrager and Culler may, in fact, be of type-T rather than of type-L. Such a possibility is supported by the fact that these investigations conditioned a rapid twitch of the semitendinosus muscle instead of a complete flexing movement of the limb, and that they also used the method of shock-shock conditioning.

(Manuscript received March 7, 1942)

References

1. Headlee, C. R., & Kellogg, W. N. Conditioning and retention under hypnotic doses of nembutal. *Amer. J. Psychol.*, 1941, **54**, 353–366.
2. Kellogg, W. N., Davis, R. C., & Scott, V. B. Refinements in technique for the conditioning of motor reflexes in dogs. *J. exp. Psychol.*, 1939, **24**, 318–331.
3. Shurrager, P. S., & Culler, E. Conditioning in the spinal dog. *J. exp. Psychol.*, 1940, **26**, 133–159.
4. Wendt, G. R. An analytical study of the conditioned knee-jerk. *Arch. Psychol.*, 1930, No. 123, 97 pp.

METHODOLOGICAL EVALUATION OF THE GALVANIC SKIN RESPONSE, WITH SPECIAL REFERENCE TO THE FORMULA FOR R.Q. (RECOVERY QUOTIENT)

BY G. L. FREEMAN AND E. T. KATZOFF

Northwestern University

The increasing importance of the galvanic skin response to the study of autonomically (**4**) and centrally (**11**) controlled activities of the intact organism has led several investigators to give serious consideration to the quantitative expression of the relations between various aspects of the phenomenon. As far back as 1913, Dodge (**5**) pointed out that the dynamics of a physiological response curve (such as the stimulus-produced changes in blood pressure or galvanic skin conductance) was necessarily a function of preëxisting conditions of the organism. Subsequently Darrow and Heath (**14**) found rate of recovery in skin conductance significantly related to the level which preceded the displacing stimulus. The first approach to the *mathematical* expression of this relationship was made by Darrow (**3**) in a study of the galvanic responses to a series of physically equivalent shock stimuli. This report indicated that both the increment and decrement phases of galvanic response curves are related to the preëxisting level of physiological activity. That is, greater excitation and recovery changes are occasioned by the stimulus when the level of background activity on which it is imposed is low. Darrow comments on these relationships as follows (**3**, pp. 292–293):

> "Assuming that the curve of rest is asymptotic to a level of equilibrium maintainable by the more persisting external and internal conditions, we should expect the greater the degree of excitation or of displacement of the galvanic curve occasioned by momentary stimulating conditions, the more rapid would be the return toward a lower level on the cessation of the stimulus. It is a common observation in the case of natural phenomena of change that the rate of approach toward an equilibrium tends to approximate proportionality with the degree of displacement from equilibrium. . . ."

Since it is manifestly impossible to obtain with the intact organism a basal excitation level free from the residual effects of preceding stimuli, and hence characteristic of the persisting internal condition of rest, Darrow's primary interest was the estimation of that level from a comparison of the tangents at two points on the 'curve of relaxation' following momentarily produced experimental displacements. He derived an elaborate set of formulæ based on log con-

ductance scores, which rationalized the exponential function of the curve of rest. The theoretical aspect of the relation between momentary response and excitation background was developed contemporarily by Freeman (**8**) in a paper on the postural substrate, and the relation between characteristics of the momentary 'startle response' curve and the change in conductance level produced by sustained stimulation was included in a later report by the same author (**7**). Wenger and Irwin (**15**) contributed to a related problem, and concluded that the true correlation between resistance level and magnitude of ohmic change is unity.

The purpose of the present report is to determine the best mathematical expression of the ratio between the excitation and recovery phases of galvanic response curves. To this ratio Freeman (**8**) has already applied the term *R.Q.*, or Recovery Quotient, and has published preliminary data indicating its significance as an index of emotionality.

On Units of Galvanic Measurement.—Measures of the galvanic skin response may be given any one of several expressions. The customary apparatus employed in galvanic measurement is a Wheatstone bridge which records in terms of ohmic resistance. Such values have often been used without translation (Landis, **12**), or in terms of standard scores (Wenger and Irwin, **15**). However, the use of units of conductance rather than resistance is to be preferred because of the more direct relation to the major physiological changes (sweating) responsible for the galvanic phenomenon (Darrow, **2**). Resistance is easily translated into conductance terms by the following formula:

$$C = 10^6 \times \frac{1}{R}$$

where C = micromho units of conductance, and R = ohms resistance. An alternative conductance unit which has been suggested by Darrow is log conductance. The mathematical expression of the relationship of this measure to changes in resistance units is as follows:

When $\frac{\Delta R}{R} = K$, then

$$100\left[\log\frac{10^6}{R} - \log\left(\frac{10^6}{\Delta R + R}\right)\right] = 100\log\left(\frac{10^6}{R}\cdot\frac{R+\Delta R}{10^6}\right)$$

$$= 100\log\frac{R+\Delta R}{R} = 100\log\left(1+\frac{\Delta R}{R}\right) = 100\log(1+K) = k.$$

Empirical verification is provided by computing the absolute difference in log conductance for a fixed percent change in resistance.

Thus, if skin resistance changes from 20,000 to 22,000 ohms, the percent change in resistance is equivalent to an absolute change in log conductance of 4.1;[1] a similar change in resistance from 30,000 to 33,000 ohms will also equal an absolute change in log conductance of 4.1. This equivalence of log conductance for a given percent change in resistance will hold throughout the entire range. However, because of the reciprocal relation between units of physiological activity (amount of sweating) and skin resistance units, a physiological change indicated by a shift from 20,000 to 22,000 ohms is not equivalent to that indicated by a shift in resistance from 30,000 to 33,000 ohms.

If we accept Darrow's evidence that conductance varies directly with the physiological activity (palmar sweating) basic to it, it follows that an absolute change in micromhos conductance is equivalent to an absolute change in physiological activity, and is, consequently, the proper measure of that physiological change. Thus, if we wish to study the relative effects of a standardized stimulus when it is superposed on different activity levels (excitation backgrounds) we should compare absolute changes in conductance. If, however, our purpose is to measure the effects of the stimulus on the organism independent of the base upon which it is superimposed, then we should use percent changes in conductance.

The theoretical justification for this point of view is that the increase in activity due to a standardized stimulus would be greatest when imposed in the middle range of 'basal' activity and least at the extremes of this range. Thus, an individual who is 'relaxed' to a point of drowsiness or 'mobilized' to a state bordering on emotional collapse will not react as extensively to a stimulus as when in a state of normal attentiveness. The usual level of background activity upon which experimental stimuli are imposed varies from the lower to the middle of the range (*i.e.*, subjects are seldom placed in so critical a state as to become completely upset).

Darrow has suggested that the log of the conductance should be used to measure stimulus effects at different levels of physiological activity. As far as we can see, the only advantage to the employment of log conductance over percent change in conductance is that the log measure spreads the scores over a greater range; thus a resistance range of 90,000 ohms is equivalent to a log conductance range of 100 units and to a percent conductance range of 90 units (see Fig. 1). In accomplishing this advantage, however, Darrow's unit (100 log conductance) involves a superfluous mathematical manipulation (*i.e.*,

[1] A conversion table for changing resistance scores to conductance and log conductance scores has been prepared by Dr. L. H. Sharp and may be obtained by addressing the Laboratory of Psychophysiology, Northwestern University.

it takes the exponential function of a reciprocal function which is already a direct measure of the physiological activity involved. This is equivalent to the use of the log of the basal rate of oxygen consumption as a measure of metabolism.[2]

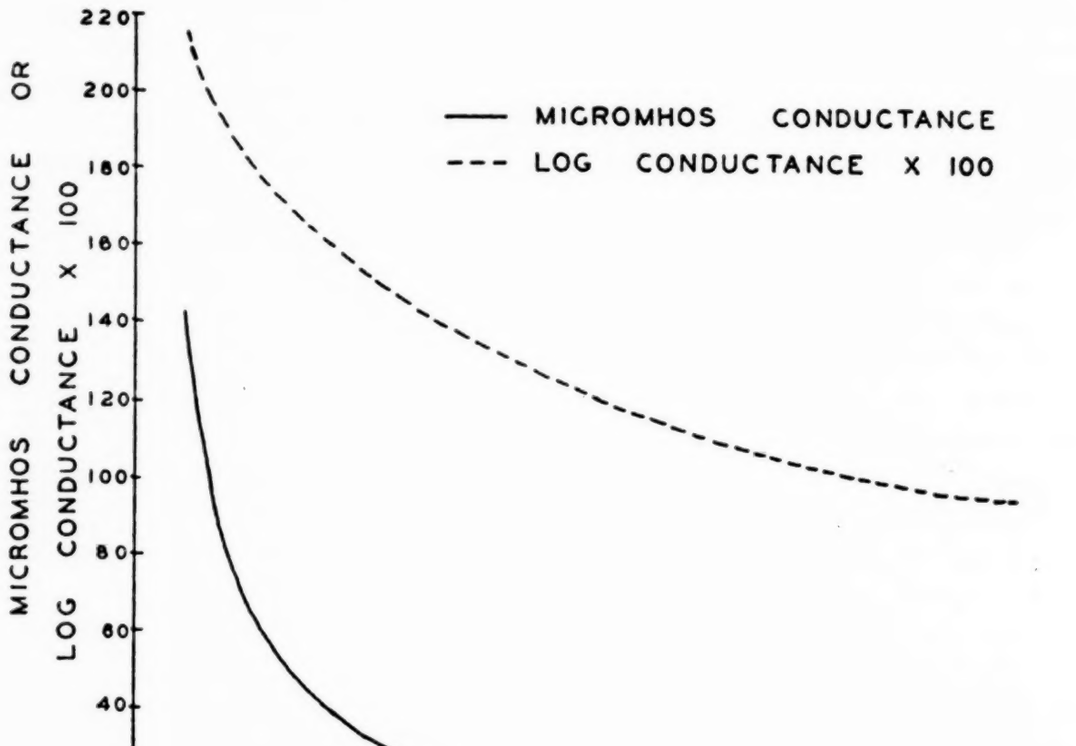


FIG. 1. Relation of ohmic resistance to derived measures of physiological activity.

On the Derivation of the R.Q. Formula.—Before considering the derivation of measures which will express the ratio of 'arousal' to 'recovery' in galvanic response curves, the theoretical significance of the concept of the recovery quotient (R.Q.) requires discussion.

[2] Of course, at the present state of our knowledge, one might argue for use not only of percent and log conductance scores, but also for percent ohmic resistance, and standard scores as well. The superiority of conductance units rests upon the exactitude with which micromhos parallel drops of sweat, and this needs to be checked with more cases. One should not neglect the physical error in making measurements. Size of electrode and amount of current used make differences which should somehow be equated; but even with these constant, differences in thickness of the skin, etc., might also have to be equated before galvanic changes could be used critically to compare physiological reactivity from individual to individual. Until more research of a strictly physiological nature is done, it will probably be impossible to settle on any final 'best' unit of measurement.

In dealing with palmar skin conductance changes (either the 'reflex' effects of a single startle stimulus, or the more prolonged changes in level resulting from a series of stimuli or a protracted task), we are concerned with a response which has been fairly well established as an index of the general activity level of the total organisms (**9**). When *S* is placed in a resting condition, palmar skin conductance usually drops to a rather stable (basal) level. If *S* is then stimulated, the conductance rises sharply. When stimulation ceases, conductance begins to return toward the level from which the original change occurred. While both the rise and fall of this curve are indicative of change in the general level of activity, we may consider one position an index of 'arousal' or 'mobilization,' and the other position as an index of 'recovery.' The theoretical justification for this approach lies in the well-recognized principle of homeostasis (**1**), or the tendency of an organic system to react to a displacing stimulus in such a way as to remove the stimulus or its effects and so restore the equilibrium of a 'steady state.'

If we think of the total galvanic curve as expressing a homeostatic process, it follows that the peak of the curve is the transition point at which response begins to overcome the displacing effects of stimulation (*i.e.*, restorative processes exceed the displacing properties of the stimulus). Thus, if an organism in a stable resting state is exposed to a standard stimulus, the percent change in activity from that level to the peak level is an index of the displacing effect of that stimulus upon the organism. Furthermore, the time taken for that level of activity to return from a given peak to the pre-stimulus base is an index of the efficiency of the organic equilibratory processes (discharge of stimulus-aroused energies). It is to be noted that it frequently takes the organism fifteen minutes or longer to discharge the effects of experimental stimulation sufficiently to produce a reasonable return of skin conductance to the pre-stimulus base. Since it is exceedingly difficult to isolate the *S* from uncontrolled stimulation during so long a period, the usefulness of a simple time-ratio between the 'arousal' and 'recovery' parts of the galvanic response curve is limited.

A more promising expression of this same relationship is the ratio between the arousal increment which obtains at some standard time after beginning of stimulation and the recovery decrement which obtains at some stipulated time after the peak of the response curve has been reached.[3] Results from this laboratory (**12**) indicate

[3] The standard intervals chosen must be appropriate to the experimental situation; that is, a momentary startle stimulus requires short mobilization and recovery intervals. Ideational stimuli of longer duration and effect require the use of longer mobilization and recovery intervals. The selection of appropriate time constants should be based on the modal distribution of intervals obtaining between the three points of inflection on the galvanic response curves.

a high correlation in rats between the drop in physiological response level (breathing rate) 5 minutes after the peak reaction and the length of time taken for complete recovery. In reactions of human subjects to stimuli of long duration it was found that individuals tended to reach peak reaction within a minute after cessation of stimulus and tended to maintain the same rank in 'recovery' regardless of whether the scores were based upon readings taken at 5 or 10 minutes after this 'peak.' As might be expected, the readings taken 5 minutes after peak arousal were considerably more stable than those which obtained at the 10 minute period.

It has thus been our practice to represent 'mobilization' by percent conductance increment one minute after cessation of stimulation and 'recovery' by percent conductance decrement 5 minutes after this peak. If we represent the galvanic response curve as in Fig. 2, the amount 'mobilized' can be expressed by the distance

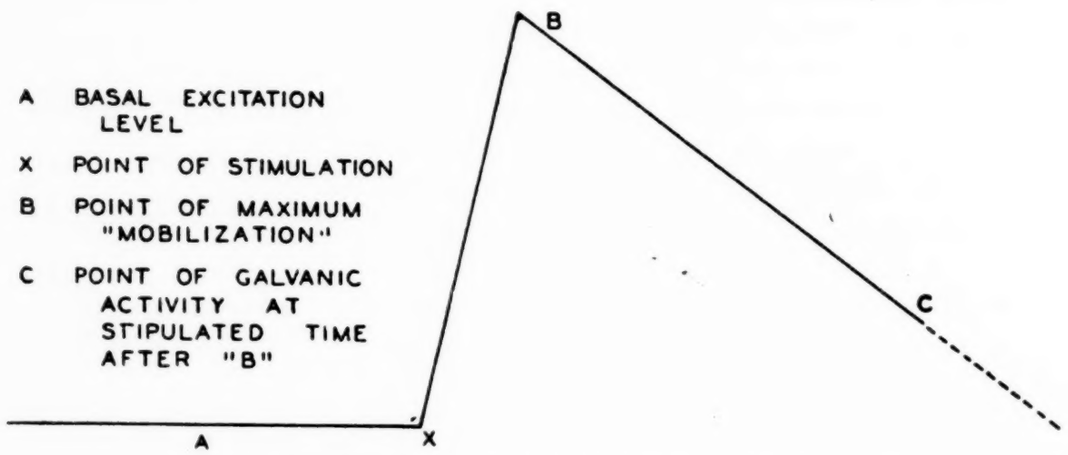


FIG. 2. Schematic representation of the galvanic response curve

A–B, while the amount recovered can be expressed as the distance *B–C*. With time relations between *A*, *B*, *C* in some constant ratio (1 : 5 in our practice), the equilibratory efficiency, or *Recovery Quotient* to a given stimulus displacement, is measured by the ratio of the mobilization and recovery changes. In terms of our schematic figure the formula is $\frac{(B - C)}{(B - A)} = \text{R.Q.}$ By using conductance changes for $(B - C)$ and $(B - A)$ the Weber-Fechner function is still included even though the base (A) cancels out in the formula.[4]

[4] H. Jost, in an unpublished thesis (Univ. of Chicago, 1940), has expressed the 'galvanic equation' in terms of the *resistance* change. Since our measures of physiological activity are conductance changes based upon a resting conductance value, we have already taken into account the level from which such changes occur. Jost's formulation is in error due to its failure to take into account this aspect of the Weber-Fechner function.

On the Stability of the R.Q. Measure.—Although it was by way of mathematical operations that the first hint appeared of the R.Q.'s independence of its basal level, it was not long before experimental evidence added weight to this contention. A reëxamination of some data previously reported (**9**) showed that the R.Q.'s to a standard displacing stimulus given to the same subjects before and after the release of micturational tension did not differ significantly. The correspondence in R.Q. (computed either in log conductance or micromho conductance units) obtained despite the fact that basal level before micturition was significantly higher than the post-micturitional basal. As can be seen in the summary of these data (Table I) the use of log conductance conceals the true difference in basal measures.

TABLE I

SUMMARY OF DATA ON MOBILIZATION TO A CONSTANT STIMULUS FROM LOW AND HIGH 'BASAL' TENSION CONDITIONS

	Mean Basal		Mean Mobilization		Mean R.Q.	
	Cond.	Log C.	Cond.	Log C.	Cond.	Log C.
High tension condition	163.5	1210.6	172.7	1235.2	1.32	1.40
Low tension condition	153.7	1180.9	168.4	1222.7	1.16	1.23
Values of *t* for differences between high and low basal tension conditions	4.42	1.44	1.54	.7418	.31	.28

A schematic representation of the relationship of galvanic response curves to equivalent stimuli at different basal levels is represented in Fig. 3. Here the straight line *XY* represents possible variation in basal level, while the exponential curve *BB* represents the peak level to which *S* will mobilize from various 'basals' in response to a standard physical stimulus. The dotted triangles (*ABC*), which are a representation of individual response curves from different bases, are seen as similar. The ratio of the ascending to the descending changes in skin conductance is shown as being constant throughout.

The empirical data upon which this schematic representation is based are plotted as dots about the exponential function. These points were obtained by taking the difference between base and peak conductance for equivalent stimuli at different basal levels for *20* cases.

Comment.—The study of galvanic skin responses has now reached the stage where exact quantitative expression is required in order to have significance for behavioral research. The day is past when the psychologist may attach a galvanometer to the subject to determine whether he is experiencing an 'emotion,' and in its place there is a

growing recognition that the galvanic response curve may be assayed as the index of total behavior dynamics without regard to the specific pattern of response, the mental state, or the stimulus which was utilized. Regardless of the level of organic activity which is met by the stimulus, stable dynamic relations appear to obtain between the several variables involved.

The formulation of the laws governing the dynamics of the galvanic response curve is still in a nebulous state. But unless we

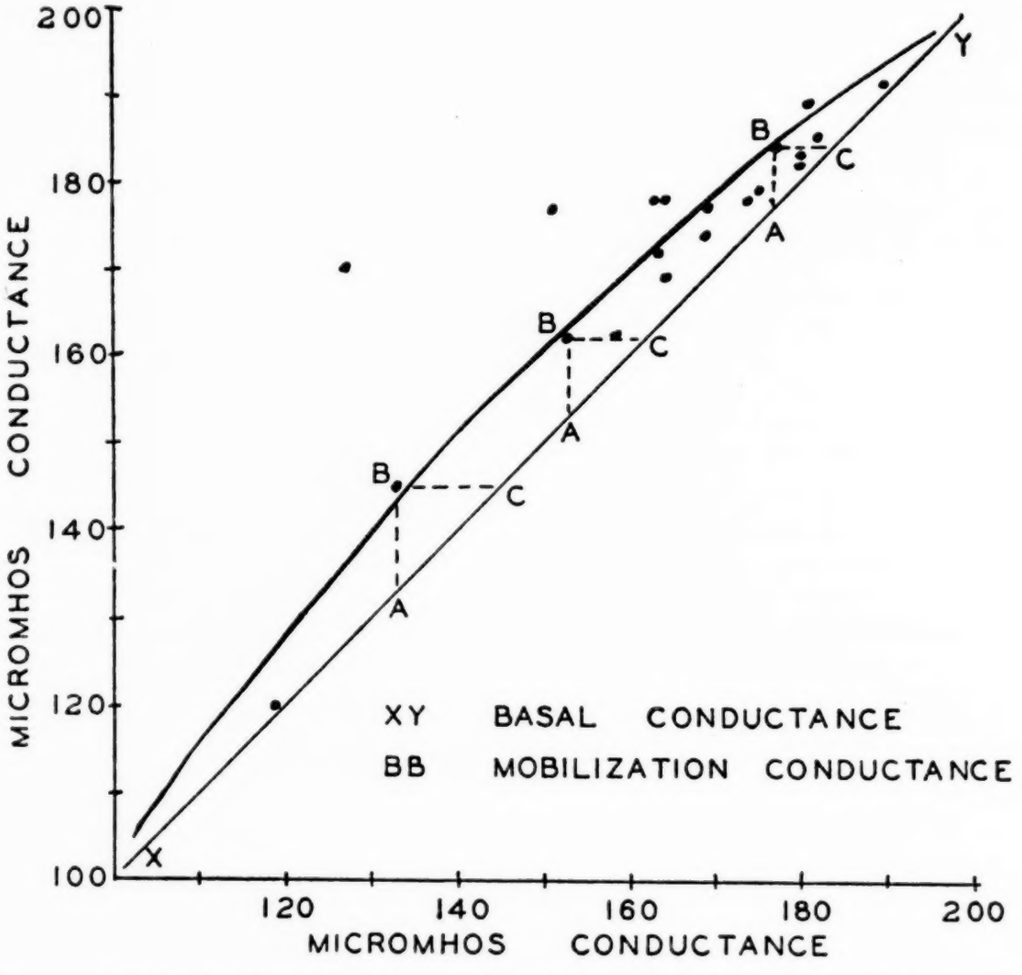


FIG. 3. Relation between 'basal' conductance (*XY*) and 'mobilization' conductance (*BB*).

recognize that such laws exist, and that all organisms are subject to them, a separate science will be required for each organism and every situation.

Methodological papers are rare in psychology, and this probably accounts for the vast number of unrelated and unfruitful studies utilizing physiological measures. The study of galvanic responses is a case in point. Investigators have seldom sought to integrate the various aspects of the phenomena, but have taken some isolated

attribute such as the amount or latency of the response and correlated it with other static measures such as performance score, body weight, or verbally named emotion. We have already discussed developments from the first integration of the galvanic response curve by Darrow and Heath down to the present exposition of conductance R.Q. It is expected that subsequent investigations will contribute further to this type of dynamic analysis.[5]

The question arises as to what extent one may generalize the concept of Recovery Quotient to measures other than the galvanic response. It is obvious that part-mechanisms unless directly involved in a response situation will not yield a typical equilibratory curve. For example, the peristaltic movements of the intestinal tract will probably not be significantly affected by the eyelid reaction to a puff of air. Hence it would be futile to estimate R.Q.'s for airpuff reactions from fluoroscopic registrations of intestinal movements. It is profitable to record instead the activity of more directly involved organ systems, such as associated skeletal muscles and verbal reactions. The relation between these specific reactions and the more general (galvanic) response curves holds promise of being indicative of the ratio of overt reaction to general recovery. Any part-response which is significantly involved in a displacing situation can be made the basis for an R.Q. determination, and positive correlations have been found to exist among the R.Q. measures computed from concurrent reactions. A previous report (**9**) has shown positive correlations between galvanic, blood pressure and electromyrographic R.Q.'s, and has further indicated a method by which the R.Q. concept may be applied to performance-test scores. The intercorrelations of R.Q. measures obtained on both covert and overt aspects of total response suggest a certain validity in generalizing from one appropriate measure to another.

Basic to the entire concept of arousal and recovery is Cannon's theory of homeostasis (**1**). By this doctrine is meant the tendency of the organism to maintain constant states by reacting to displacement from such states in such a manner as to reëstablish them. Cannon has demonstrated the operation of homeostatic processes in limited organ systems, and in maintenance of states integrated through the autonomic nervous system. This doctrine is equally applicable to behavior integrated mainly through the central nervous system. In such neuromuscular homeostatic reactions, the organism is usually displaced (aroused) by external stimuli, and behaves in a manner ultimately resulting in reëquilibration (recovery). In this

[5] E. F. Dunbar (**6**) has indicated the need for developing from static physiological measures dynamic derivatives of the "first and second order." Only such measures, according to Dunbar, may be expected to yield significant correlations with psychosomatic disorders.

sense, the galvanic R.Q. may be considered a measure of the homeostatic efficiency of the organism.

(Manuscript received January 27, 1942)

References

1. Cannon, W. B. *The wisdom of the body.* New York: W. W. Norton & Co., 1940.
2. Darrow, C. W. The significance of skin resistance in the light of its relation to the amount of perspiration. *J. gen. Psychol.*, 1934, **11**, 445–452.
3. Darrow, C. W. The equation of the galvanic skin reflex curve. *J. gen. Psychol.*, 1937, **16**, 285–309.
4. Darrow, C. W., & Darling, R. Determining activity of the autonomic nervous system from measurement of autonomic change. *J. Psychol.*, 1938, **5**, 85–89.
5. Dodge, R. Mental work, a study of psychodynamics. *Psychol. Rev.*, 1913, **20**, 1–40.
6. Dunbar, H. F. *Emotions and bodily changes.* (2nd ed.) New York: Columbia Univ. Press, 1938.
7. Freeman, G. L. The effect of inhibited micturition upon interrupted and completed acts of unrelated origin. *J. gen. Psychol.*, 1938, **19**, 277–283.
8. Freeman, G. L. The postural substrate. *Psychol. Rev.*, 1938, **45**, 324–334.
9. Freeman, G. L. Postural accompaniments of voluntary inhibition of micturition. *J. exp. Psychol.*, 1938, **23**, 45–61.
10. Freeman, G. L., & Griffin, L. L. The measurement of general reactivity under basal conditions. *J. gen. Psychol.*, 1939, **21**, 63–72.
11. Freeman, G. L. Toward a psychiatric plimsoll mark: physiological recovery quotients in experimentally induced frustrations. *J. Psychol.*, 1939, **8**, 274–292.
12. Krasno, L. R. Effect of displacing stimuli on the subsequent respiratory activity of rats and guinea pigs. Unpublished Thesis, Northwestern Univ. Library, Evanston, Ill., 1939.
13. Landis, C., & DeWick, H. N. The electrical phenomenon of the skin (psychogalvanic reflex). *Psychol. Bull.*, 1932, **29**, 693–752. See also pages 185 and 527, same journal, 1929.
14. Lashley, K. S. (Ed.) *Dynamics of behavior.* Chicago: Univ. Chicago Press, 1932. See report of Darrow & Heath.
15. Wenger, M. A., & Irwin, O. C. Fluctuations in skin resistance of infants and adults and their relation to muscular processes. *Univ. Ia. Stud. Child Welf.*, 1936, **12**, 143–178.

GALVANIC SKIN RESPONSES OF INFANT CHIMPANZEES [1]

BY AUSTIN H. RIESEN

Yale Laboratories of Primate Biology

The purpose of this study was to determine the presence of the galvanic skin response with infant chimpanzees and to investigate its correlates. Of particular interest was the question of sleeping *vs.* waking resistance levels. Results with human infants have heretofore been equivocal (**5, 6, 7, 8**) on this problem. The six bottle-raised infants of the experimental nursery at the Yale Laboratories of Primate Biology which were used as subjects ranged in age from 12 hours to 36 weeks at the time of the testing.

METHODS

Apparent skin resistance was recorded by means of a modified Davis (1) vacuum-tube circuit, which maintained a constant current through the subject. The value of the imposed current used was 40 microamperes. Hard rubber electrode cups, having a zinc plate glued to the bottom with collodion, were filled with an electrolyte of 1 percent zinc sulphate stiffened to a gel with 8 percent agar-agar. The inside dimensions of the cups were ½ cm in depth and 2 cm in diameter.[2] The electrodes were fastened with adhesive tape to the plantar surface of the right foot, the active electrode area, and to the right lower leg where the skin had first been pierced with a sterile surgeon's needle to form the indifferent electrode area. The subject was left unrestrained in the crib except for the loose electrode leads on the leg.

The first series of measurements was made to investigate the relation of apparent skin resistance to changes in the general level of activity of the organism. During continuous sleep, readings were taken at 5-minute intervals. More frequent readings were taken during active states and during changes from one state to another. The activity of the infant, the values of voltage and current in the amplifier, and the time constituted the data recorded.

A second series of observations was made in conjunction with the application of tactual-kinæsthetic, sound, and visual stimuli.

RESULTS

Skin Resistance and Activity Level.—Results obtained with the youngest and oldest subjects are shown in Figs. 1 and 2, respectively. Jed had been asleep several minutes prior to the start of his record at 9:00 P.M., which was 12 hours after birth. The double record constitutes a validation of the effectiveness of the technique for eliminating reactivity at the 'indifferent' electrode. With this

[1] Publication of the Infant Studies Program, supported in part by The Samuel S. Fels Fund. Data presented in part before the Florida Academy of Sciences, Florida Southern College, Nov. 21, 1941.

[2] To provide comparability, electrodes and current value employed were duplicated from those now in standard use with children by Dr. M. A. Wenger at The Samuel S. Fels Research Institute, Antioch College, Yellow Springs, Ohio.

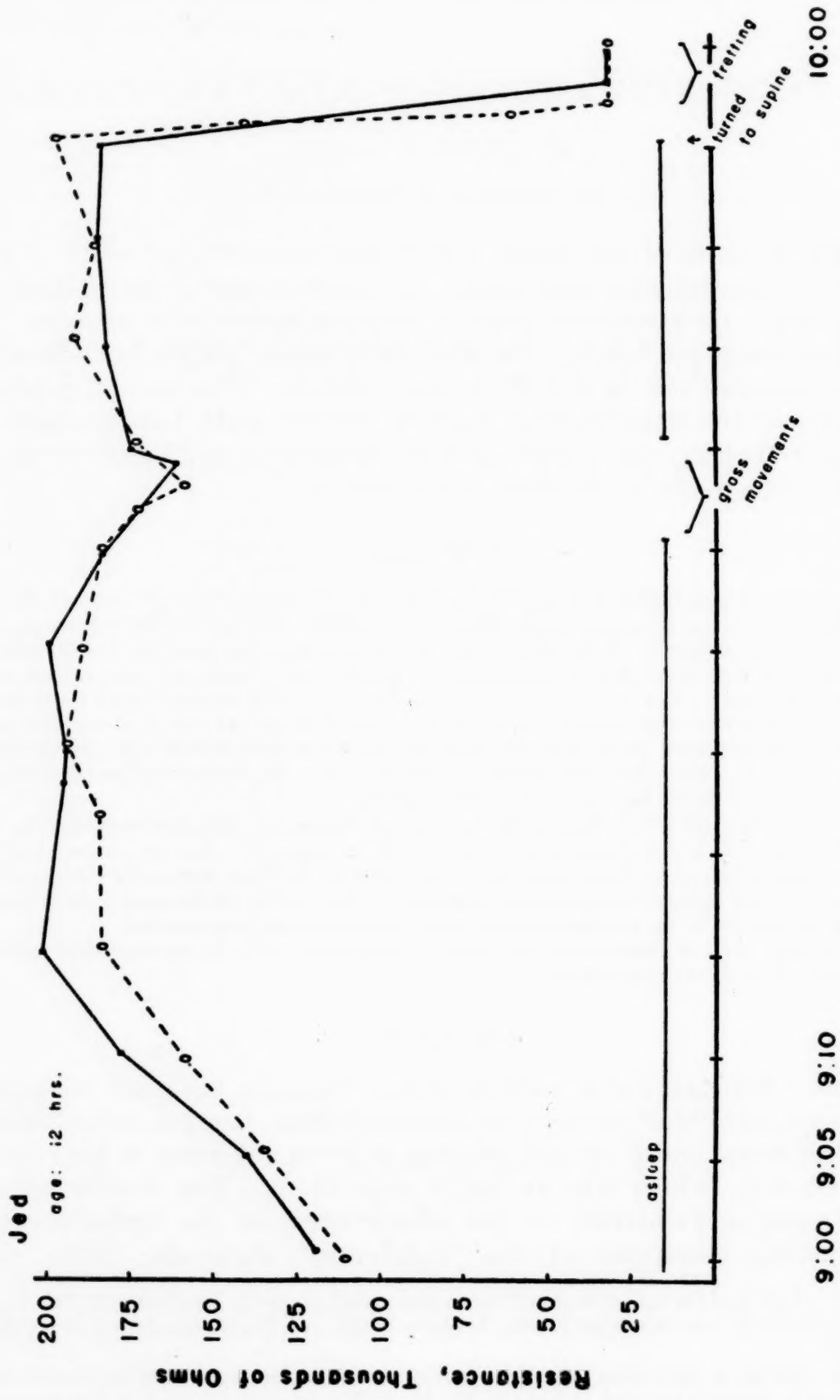


FIG. 1. Apparent plantar skin resistance of a male infant during the thirteenth hour after birth.

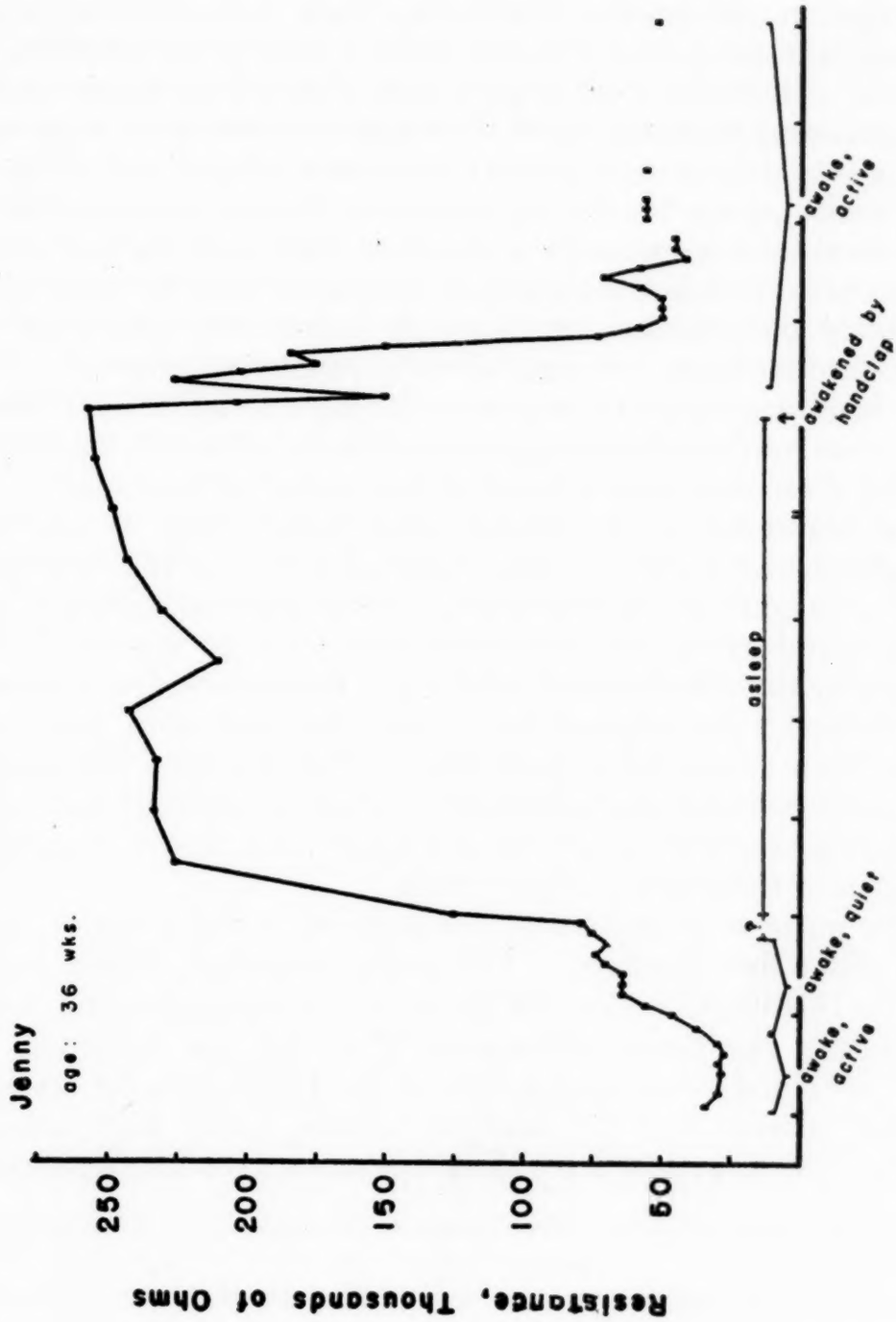


FIG. 2. Fluctuations in apparent skin resistance of a female of 36 weeks during active and sleeping states.

subject a third electrode was taped to the left plantar skin surface, after this area had also been pricked. The dashed line connects values of resistance obtained when this electrode was switched in in place of the one on the right lower leg. Unless piercing the skin was effective in eliminating reactivity, these two different areas should each have summated with the active area to give very different total values. Since the method gave only closely contiguous rather than simultaneous readings, small discrepancies were to be expected.

From a moderately high plantar resistance at 9:00 P.M. (Fig. 1) the level continues to rise for 15 minutes. During continuation of sleep this level then remains fairly constant until near the end of the hour, with the exception of a dip at 9:35 associated with the occurrence of gross body movements. At 9:56 the infant was awakened by turning it onto its back. A rapid drop in resistance followed. The resistance level remained at a minimum during a period of fretfulness. Readings were continued through a second hour; although not shown graphically, these data are included in the statistical treatment.

At the beginning of the session with Jenny (Fig. 2) she was slightly apprehensive and the skin resistance was at the relatively low level of 25,000 to 30,000 ohms. After gradual quieting the infant fell sound asleep and remained asleep for 50 minutes. She was aroused by the noise of hand clapping. Resistance level returned to approximately the original low value, but not until after the infant had been active for a short time. This lag upon awakening finds its converse in the gradualness of increase in apparent resistance as the infant passes from an active to a quiet state, and is characteristic of the data from each subject tested.

For the purpose of statistical comparison, activity states were classified under five headings. The mean resistance values corresponding to this classification are given in the accompanying tabulation. Highly significant differences ($P < .01$) are indicated by values of t obtained in the comparison of the Jenny data for 'awake, active' with 'asleep' ($t = 9.5$) and for 'awake, quiet' with 'asleep' ($t = 12.4$). (*Cf.* **3**, p. 128 ff.) More adequate data for the session

	Jenny		Jed	
	N	Mean	*N*	Mean
Fretting	0		5	30.6
Awake, active	26	67.9	9	54.0
Awake, quiet	8	58.9	3	65.0
Asleep, doubtful	3	74.3	10	160.2
Asleep	11	226.6	33	150.7

with Jed provide a comparison of 'fretting' with 'awake, active' which gives a t of 4.4, still well above the value of 3.06 required for an n of 12 to indicate a probability of 1 in 100 that the difference occurred by chance. The t-value for 'awake, active'—'asleep, doubtful' is 20.2; for 'awake, active'—'asleep,' 6.7; and for 'fretting'—'asleep,' 12.6. The two inversions in the order of increasing mean resistances found in the tabulation can be accounted for in terms of the lag phenomenon mentioned in the paragraph above. In addition, inaccuracies of judgments, especially with respect to sleeping states, attenuate the data. Four other infants, whose ages were 1 day, 2 weeks, 9 weeks, and 15 weeks, respectively, were used in this phase of the experiment. All gave reliably higher mean resistances during sleeping states as compared with active waking states. The results on this broader aspect of the problem are conclusive.

Wenger and Irwin (**8**) found a close correspondence in adult human subjects between changes in palmar and plantar skin resistance and degree of muscular activity. They interpret their results with human infants, similar to those described above for chimpanzees, as reflecting this correspondence. The apparent failure of Landis (**6**) to find changes in resistance associated with sleeping and active states is explained by Wenger and Irwin in terms of the technique employed and the consequent low values of apparent resistance he obtained. Freeman and Darrow (**4**) report that: "Instances in which adult human subjects were judged by their breathing and closed eyes to be asleep were uniformly characterized on waking by a marked increase in palm conductance." Richter's (**7**) suggestion that apparent skin resistance provides a criterion of sleep can be accepted as valid only with qualifications. Chief among these would appear to be that lag in resistance values produces inversions in the order of these values in connection with changes from sleep to waking and waking to sleep.

Resistance During Trance-Like State.—The infant, Flora, at 9 weeks of age, and again during an experimental session a week later, assumed a posture not consonant with the relaxation ordinarily associated with sleep, but during which the resistance level increased to within the range of values obtained with her during sleep. Her head and upper chest were raised, their weight supported on the forearms as she lay prone. Her face was toward the closed corner of the crib, and her eyes were open, turned upward toward the ceiling of the room. The visual field thus afforded was highly homogeneous, being somewhat brighter above because of a more strongly illuminated ceiling, but a light ivory color throughout. The subject engaged in a slow rhythmic swaying motion from side to side during much of the time that she held this position. The

readings were first classified under the 'awake, active' category, but it soon became obvious that they did not belong to the same distribution. When classified separately, the means of the 'active' and 'trance' categories were found to be significantly different for each of the two sessions. The tabulation below gives the number of readings and mean values for various activity states during the session on each day. Values of t for the differences between the means of 'awake, active' and 'trance' are 3.15 ($P < .01$) for April 4 and 5.18 ($P < .01$) for April 10. The mean resistance during the

	April 4		April 10	
	N	Mean	*N*	Mean
Awake, active	12	96.2	11	75.7
Asleep, doubtful	15	109.4	5	97.8
'Trance'	11	123.3	8	125.6
Asleep	10	136.5	3	133.9

'trance' state is significantly higher than during 'asleep, doubtful' for April 10 ($t = 3.26$; $P < .01$), but this comparison does not prove significant for April 4 ($t = 1.21$; $P < 0.3$), on which day the 'asleep, doubtful' readings were highly variable.

The high resistance found for the trance-like state recalls the conclusion of Davis and Kantor (**2**) that "the lethargic condition of hypnosis resembles sleep in the behavior of skin resistance; the active hypnotic condition resembles the waking state." The lethargic condition, these authors state, may be one into which the subject slips when no stimulation is given during the hypnotic state.

Since considerable tonus in certain muscle groups was required for maintaining the posture described above, the data would seem to limit the generality of the correspondence between resistance level and degree of muscular tension. It is possible that processes independent of, but still highly correlated with, muscular tonus are more closely reflected by level of skin resistance than is muscular tension itself.

Galvanic Changes Following Sensory Stimulation.—In human subjects the typical galvanic skin response to a sudden stimulus has a latency of approximately two seconds and consists of a rapid drop in resistance with a gradual return to the original level. Jones (**5**) has demonstrated the presence of this response in infants as young as three months. A series of tests with sound stimuli exhibited considerable intra-individual variability of response in the chimpanzee infants, which is characteristic also of results with human infants. Thus Jed, aged 4 months, gave responses to pure tones in 56 percent

of the trials. On seven of the ten failures of response the infant was lying fully relaxed or possibly asleep. He was awake and holding his head raised during all trials which gave strong galvanic responses. Head movements occurred more often than not on these trials, indicating that the galvanic was a component of a more generalized response pattern. Ear twitching, or Preyer's reflex, occurred on two of the three trials for which the response was classified as slight, *i.e.*, under 1,000 ohms.

The female infant, Fanny, at 2 weeks of age, exhibited only four galvanic responses in ten presentations of tones of 512, 1,024, and 2,048 dv./sec. No responses occurred to the highest pitch but the data are obviously insufficient to provide evidence for differential auditory sensitivity. Again, as is typical with human infants, relaxed states were associated with reduced frequency of galvanic responses. In this younger infant, as well as with Jed, the latencies of the galvanic responses were between two and three seconds.

The most effective response elicitors used were complex tactual-kinæsthetic stimuli produced by rolling the infant or by lifting and lowering it. Resistance drops of from 7,000 to 50,000 ohms followed these stimuli. Tactual stimulation alone, on the other hand, frequently resulted in no detectable galvanic change. Pinching the lower leg produced resistance decreases of 6,000 to 12,000 ohms in subjects of 2 weeks and 4 months of age.

Visual stimuli failed to produce resistance changes in a two-weeks-old infant. The tests were not exhaustive. Infants of 2½ and 4 months of age responded with pronounced decreases to the sudden introduction of large objects into the visual field. Gross muscular startle responses usually accompanied the galvanic responses, which again were about two seconds in latency.

Conclusions.—The relational aspects of the galvanic skin response are essentially the same in infant chimpanzee as in man. This finding carries with it the implication that the galvanic skin response provides a means for detecting autonomic nervous system function in the chimpanzee as it does in man (**9**).

Summary

Significantly higher mean values of apparent skin resistance during sleep as compared with active waking states were found in each of six infant chimpanzees tested. With a shift from one state to another there was a lag in the resistance change followed by gradual stabilization of the ohmic value. One infant spontaneously assumed a trance-like state during which tonus was apparently maintained at a fairly high level in certain muscle groups. Mean

resistance values during these periods lay between the means for 'asleep' and 'asleep, doubtful' categories. Implications of the data for physiological correlates of the galvanic skin response are suggested.

Responses to certain sensory stimuli exhibited characteristics with chimpanzees similar to those previously found for human subjects.

(Manuscript received March 11, 1942)

References

1. Davis, R. C. A vacuum tube for stabilizing the current during measurements of the galvanic reflex. *Amer. J. Psychol.*, 1929, 41, 474.
2. Davis, R. C., & Kantor, J. R. Skin resistance during hypnotic states. *J. gen. Psychol.*, 1935, 13, 62–81.
3. Fisher, R. A. *Statistical methods for research workers.* London: Oliver and Boyd, 1938.
4. Freeman, G. L., & Darrow, C. W. Insensible perspiration and the galvanic skin reflex. *Amer. J. Physiol.*, 1935, 111, 55–63.
5. Jones, H. E. The galvanic skin reflex in infancy. *Child Develop.*, 1930, 1, 106–110.
6. Landis, C. Electrical phenomena of body during sleep. *Amer. J. Physiol.*, 1927, 81, 6–19.
7. Richter, C. P. High electrical skin resistance of new-born infants and its significance. *Amer. J. dis. Child.*, 1930, 40, 18–26.
8. Wenger, M. A., & Irwin, O. C. Fluctuations in skin resistance of infants and adults and their relation to muscular processes. *Univ. Iowa Stud. Child Welf.*, 1936, 12, 141–179.
9. Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1938.